



Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Comparative transcriptome analysis of chemosensory genes in two sister leaf beetles provides insights into chemosensory speciation



Bin Zhang^a, Wei Zhang^b, Rui-E Nie^a, Wen-Zhu Li^a, Kari A. Segraves^{c,d,*},
Xing-Ke Yang^{a,**}, Huai-Jun Xue^{a,***}

^a Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

^b State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

^c Department of Biology, Syracuse University, 107 College Place, Syracuse, NY 13244, United States

^d Archbold Biological Station, 123 Main Drive, Venus, FL 33960, United States

ARTICLE INFO

Article history:

Received 4 June 2016

Received in revised form

1 November 2016

Accepted 6 November 2016

Available online 9 November 2016

Keywords:

Antennal transcriptomics

Chemosensory genes

Chemosensory speciation

Chrysomelidae

Host plant use

Positive selection

Pyrrhalta

ABSTRACT

Divergence in chemosensory traits has been posited as an important component of chemosensory speciation in insects. In particular, chemosensory genes expressed in the peripheral sensory neurons are likely to influence insect behaviors such as preference for food, oviposition sites, and mates. Despite their key role in insect behavior and potentially speciation, the underlying genetic basis for divergence in chemosensory traits remains largely unexplored. One way to ascertain the role of chemosensory genes in speciation is to make comparisons of these genes across closely related species to detect the genetic signatures of divergence. Here, we used high throughput transcriptome analysis to compare chemosensory genes of the sister leaf beetles species *Pyrrhalta maculicollis* and *P. aenescens*, whose sexual isolation and host plant preference are mediated by divergent chemical signals. Although there was low overall divergence between transcriptome profiles, there were a number of genes that were differentially expressed between the species. Furthermore, we also detected two chemosensory genes under positive selection, one of which that was also differentially expressed between the species, suggesting a possible role for these genes in chemical-based premating reproductive isolation and host use. Combined with the available chemical and ecological work in this system, further studies of the divergent chemosensory genes presented here will provide insight into the process of chemosensory speciation among *Pyrrhalta* beetles.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Plant-feeding insects are an exceedingly species rich group. Part of this diversity may be explained by increased speciation within lineages as they evolve mechanisms to facilitate host plant use (Ehrlich and Raven, 1964). Specialization to feed on particular host plants can fuel adaptive divergence, creating host-specific populations that find, mate, and feed on different hosts (e.g., Funk, 1998; Janz et al., 2006; Stireman et al., 2005). These adaptations can require fine-tuning of the traits that allow insects to find and

select their host plants. For example, divergent selection (natural selection that results in different trait values between populations) acting on chemosensory traits can alter feeding, oviposition, and mating behaviors that can subsequently lead to reproductive isolation and speciation (Smadja and Butlin, 2009). Although the traits involved in chemosensory speciation (speciation caused by chemosensory traits) have been identified in diverse insect taxa such as moths, aphids, flies, beetles, bees, and walking-sticks (Löfstedt et al., 1991; Caillaud and Via, 2000; Ishii et al., 2001; Peterson et al., 2007; Vereecken et al., 2007; Nosil et al., 2007), genetic studies of the role of chemosensory speciation remain uncommon, particularly for non-model organisms (Brand et al., 2015).

Because olfaction is one of several essential elements of host and mate selection, identification of changes in chemosensory genes between diverging populations or closely related species can

* Corresponding author. Department of Biology, Syracuse University, 107 College Place, Syracuse, NY 13244, United States.

** Corresponding author.

*** Corresponding author.

E-mail addresses: ksegrave@syr.edu (K.A. Segraves), yangxk@ioz.ac.cn (X.-K. Yang), xue@ioz.ac.cn (H.-J. Xue).

provide valuable insight into the role of chemosensory adaptation in host shifts and speciation (Smadja and Butlin, 2009). At the species level and above, gene amplification and amino acid replacement events may play an important role during adaptive evolution. For example, a subset of smell and taste receptor genes underwent rapid evolution during host specialization in *Drosophila* (McBride, 2007); and in the pea aphid *Acyrtosiphon pisum*, host plant use was correlated with genetic divergence in chemosensory genes, suggesting a key role for these genes in host plant adaptation (Smadja et al., 2012). Perhaps more likely, however, the evolution of gene expression differences may serve as a main mechanism of chemosensory divergence. For instance, a comparison of chemosensory gene expression profiles of *Drosophila sechellia* and two sibling species showed significant changes in gene expression that may relate to host specialization (Shiao et al., 2015). Although assessing divergence in chemosensory genes is important for understanding speciation in phytophagous insects, annotated genomes remain scarce, thus hindering progress in identification of chemosensory genes. High throughput transcriptome analysis offers an alternative tool to identify putative chemosensory genes in non-model organisms (Oppenheim et al., 2015). A number of studies have taken advantage of this powerful approach (Andersson et al., 2013; Bengtsson et al., 2012; Grosse-Wilde et al., 2011; Gu et al., 2015; Legeai et al., 2011; Mitchell et al., 2012; Zhang et al., 2014b), but there are comparatively few that have compared chemosensory genes between sister species using different host plants. This type of analysis is necessary if we are to understand the mechanisms that have shaped divergence in peripheral olfactory recognition during speciation (Brand et al., 2015).

A good reason to focus on chemosensory genes expressed in the peripheral sensory neurons of insects is that these genes enable detection of chemical cues with diverse ecological functions (Dobritsa et al., 2003; Leal, 2013), making them key genetic factors that can impact insect behavior and host use. In insects, at least six gene families are involved in the detection of chemical signals: three receptor gene families that include odorant receptors (ORs), ionotropic receptors (IRs) and gustatory receptors (GRs); two binding protein gene families, including odorant binding proteins (OBPs) and chemosensory proteins (CSPs); and finally, the sensory neuron membrane proteins (SNMPs) (Benton et al., 2009; Kaupp, 2010; Robertson and Kent, 2009; Touhara and Vossahl, 2009; Vogt et al., 2009, 2015). The receptor gene families are usually expressed in the olfactory sensory neurons and are involved in the detection of volatile chemicals (ORs) (Carey et al., 2010; Hallem and Carlson, 2006; Sakurai et al., 2004; Stensmyr et al., 2012), contact chemicals or carbon dioxide (GRs) (Kwon et al., 2007; Vossahl and Stocker, 2007), and nitrogen-containing compounds, acids, and aromatics (IRs) (Abuín et al., 2011). In contrast, the binding protein gene families are highly abundant in the sensillar lymph of insects and usually function as carriers of hydrophobic scent molecules to the receptors (Sánchez-Gracia et al., 2009; Zhou, 2010). OBPs and CSPs, however, are sometimes also expressed in non-sensory tissues and may play a role in non-olfaction functions, such as development, molting (Wanner et al., 2005), and leg regeneration (Kitabayashi et al., 1998). Although their functions are still poorly understood, SNMPs have been shown to be crucial for the chemosensory process in some insects (Benton et al., 2007). Together, these genes allow insects to find and select hosts and mates. An important step in linking chemosensory genes to speciation is to assess how and to what extent divergence in these genes has occurred in closely related insect species.

An excellent study system in which to examine divergence in chemosensory genes is the elm-associated leaf beetles in the genus *Pyrrhalta*. *Pyrrhalta maculicollis* and *P. aenescens* (Insecta: Coleoptera: Chrysomelidae) are pests of elm trees, feeding on the foliage of

Ulmus davidiana, *U. laevis*, and *U. pumila*. These beetles have largely overlapping, sympatric ranges in China and similar emergence phenology (Nie et al., 2012), and so in natural populations, encounters between the species are likely. Despite this overlap, phylogenetic analysis of nuclear and mitochondrial data demonstrates that *P. maculicollis* and *P. aenescens* are sister species with no gene flow between them (Nie et al., 2012). Previous work examining the mechanism of divergence between these sister species indicates that host use plays a role in speciation. Although both species feed on the same host plant species, each specializes on a specific plant age. *Pyrrhalta maculicollis* prefers seedlings whereas *P. aenescens* feeds on adult trees (Zhang et al., 2015). These host associations appear to be driven by divergent preferences for age-specific leaf chemistry profiles. Zhang et al. (2015) conducted feeding and oviposition trials and observed that beetles preferred to feed and oviposit on leaf disks from their natal host plant age and on artificial leaf disks painted with leaf surface wax extracts from their natal host plant age. This behavioral work demonstrated that beetles are tuning in on plant chemistry to make oviposition and feeding choices. Furthermore, these species are also using chemosensory cues during mate recognition; there is strong sexual isolation caused by differences in insect cuticular hydrocarbon profiles (Zhang et al., 2014a). In behavior trials, males strongly preferred to mate with conspecific females and would also mate with heterospecific females if they were painted with conspecific cuticular hydrocarbons. Clearly, chemosensory traits are key in creating and maintaining reproductive isolation of *P. maculicollis* and *P. aenescens*, and these differences have likely led to changes in the underlying chemical detection pathways (Boake, 1991). Thus, divergence in chemosensory traits of *P. maculicollis* and *P. aenescens* provides an excellent opportunity to study the genetic basis of changes in peripheral olfactory recognition.

Here we use a transcriptome approach to identify members of six gene families involved in chemosensory perception in *Pyrrhalta* beetles. We screened for candidate genes underlying chemosensory speciation by characterizing orthologous chemosensory gene sets found in both species and comparing their levels of expression. We then examined the intensity and mode of selection on orthologous pairs of chemosensory genes. In addition, we constructed ML phylogenetic trees of these genes to examine the characteristics of *Pyrrhalta* chemosensory genes and their relationship to that of other insects.

2. Materials and methods

2.1. Insect collection, rearing and RNA sample preparation

We collected third instar larvae and pupae of *P. aenescens* and *P. maculicollis* from the northern section of Olympic Park (40.01°N, 116.39°E) in Beijing, China, on 3–5 June 2014. The two beetle species were reared to adulthood in the lab under typical field conditions. Larvae were fed fresh leaves of their natal host plant age: *P. aenescens* was fed leaves of adult *U. pumila* whereas *P. maculicollis* was fed leaves of *U. pumila* seedlings. Specifically, for each species, 50 beetles were placed in plastic cups (13.0 cm diameter, 9.0 cm deep) in a constant climate box held at 16:8 h LD and 25 °C. For the purpose of this study, three samples were sequenced from each species: heads of newly emerged adults, heads of sexually mature adults, and antennae of sexually mature adults. Once the beetles emerged and/or became sexually mature adults, we collected and pooled tissue in a tube for total RNA isolation for each sample. The head or antennae were removed using sterilized forceps and flash frozen in liquid nitrogen, and were then immediately transferred to a –80 °C freezer. Head samples contained ten males and ten females whereas antennal

samples had a mix of 50 of both sexes (Table S1).

Total RNA was extracted using TRIzol Reagent (Invitrogen, Life Technologies, USA) following the manufacturer's instructions, then was treated with DNase I (Invitrogen, Life Technologies, USA). We checked the purity of the samples using a Nanodrop Nano-Photometer spectrophotometer (NanoDrop products IMPLN, CA, USA), the concentration was assessed in a Qubit 2.0 Fluorometer (Life Technologies, USA), and the RNA integrity was verified using an Agilent 2100 BioAnalyzer (Agilent, USA).

2.2. Library construction and illumina sequencing

cDNA library for each sample was constructed using the RNA Library Prep Kit for Illumina according to the manufacturer's instructions (NEB, USA). Briefly, 3 µg of total RNA per sample were used to enrich poly (A) mRNA using oligo (dT) magnetic beads (Invitrogen, USA), and the mRNA was fragmented into small pieces using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer. First strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H) and second strand cDNA was synthesized using DNA polymerase I and RNaseH. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of the 3' ends of DNA fragments, the NEBNext adaptor with hairpin loop structure was ligated to prepare for hybridization. In order to select the fragments of preferential length, the libraries were purified using the AMPure XP system (Beckman Coulter, Beverly, USA). Subsequently, size-selected, adaptor-ligated cDNA was treated with 3 µl USER enzyme (NEB, USA) at 37 °C for 15 min followed by 5 min at 95 °C before PCR. PCR was performed using Phusion High-Fidelity DNA polymerase, universal PCR primers, and the Index (X) Primer. Lastly, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bio-analyzer 2100 system. The six libraries were sequenced in three lanes on the Illumina HiSeq 2500 platform (Illumina, USA) based on sequencing by synthesis with 100 bp paired-end reads (Biomarker Technologies, Beijing). The raw reads were saved as FASTQ files, and deposited in the NCBI Sequence Read Archive (*Pyrrhalta aenescens*: SRX2253604; *P. maculicollis*: SRX2253605).

2.3. De novo assembly and annotation

Prior to assembly, we obtained clean reads from the raw data by trimming adaptor sequences and removing low quality sequences ($Q < 20$) with more than 10% uncertain (*N*) bases using in-house perl scripts (Zhang et al., 2014b). These clean reads were then *de novo* assembled into unigenes using the short reads assembling program Trinity (Grabherr et al., 2011) with *min_kmer_cov* set to 2 and all other parameters set to default. To increase the transcriptome coverage, the clean reads from the three samples of each species were pooled during assembly. We then mapped reads back to transcripts (The mapped ratios were 82.15%, 81.79% and 81.12% in three *P. aenescens* samples, and 80.87%, 69.34% and 74.31% in three *P. maculicollis* samples), and these mapped reads were used in further analyses.

The transcripts over 200 bp in length were annotated using BLAST to search the sequences against the NCBI non-redundant (NR) protein database, and implementing HMMER to search the sequences against the Pfam protein database (Finn et al., 2014). Transcripts were functionally annotated as the identified protein or nucleic acid with highest sequence similarity using an E-value threshold of $1e^{-5}$ in BLAST and $1e^{-10}$ in HMMER. We predicted the coding sequences (CDS) using TransDecoder, translating them into the corresponding amino acid sequences. Gene ontology (GO) functional classification was performed using Blast2GO (Conesa

et al., 2005).

2.4. Identification of chemosensory genes

We identified the transcripts annotated as OBP, CSP, SNMP, OR, GR, and IR genes by conducting searches against the Nr and Pfam annotation databases. We then confirmed which of these were putative chemosensory genes by comparing their predicted CDS from BLASTp searches against the NCBI Nr database with an E-value threshold of 10^{-5} .

2.5. Phylogenetic analyses

To identify orthologous pairs of chemosensory genes between the two beetle species and to analyze the characteristics of *Pyrrhalta* chemosensory genes and their relationship to other insects, maximum likelihood (ML) trees for OBPs, CSPs, SNMPs, ORs, GRs, and IRs were constructed using the amino acid sequences derived from *Pyrrhalta* and published sequences of other species of beetles and model insects, including *Tribolium castaneum*, *Dendroctonus ponderosae*, *Ips typographus*, *Batocera horsfieldi*, *Monochamus alternatus*, *Anoplophora glabripennis*, *Phyllotreta striolata*, *Ambrosioma quadriimpressum* and *Drosophila melanogaster*. Published sequences were retrieved from NCBI (Dataset S1). Multiple sequence alignment was conducted using Mafft (online version 7.302) (Katoh et al., 2002, Katoh and Standley, 2013) with default setting (<http://mafft.cbrc.jp/alignment/server/>). The best-fit models for amino acid sequence evolution were selected using Prottest v3.4.2 (Darrriba et al., 2011). The resulting models (LG + I + G for OBPs and CSPs, LG + I + G + F for SNMPs, JTT + G + F for ORs, JTT + I + G + F for GRs and LG + G + F for IRs) were used to infer ML trees in RaxML v8.2.9 (Stamatakis, 2006) with 1000 non-parametric bootstrap replicates.

2.6. Selection analyses

To measure the intensity and mode of selection on the six chemosensory gene families, we estimated rates of non-synonymous (d_N) and synonymous (d_S) substitutions between the paired chemosensory orthologs of *P. aenescens* and *maculicollis*. Positive selection is suggested when the ratio d_N/d_S is greater than one, purifying selection is indicated when the ratio is less than one, and drift occurs when the ratio is equal to one (Zhang et al., 2006). Specifically, the paired orthologs of protein sequences for *P. aenescens* and *maculicollis* were aligned using the online version of Mafft, then were back-translated into nucleotide sequences using PAL2NAL (<http://www.bork.embl.de/pal2nal/>) (Suyama et al., 2006). Finally, the analyses of d_N/d_S were performed using model M1 in codeml of PAML package version 4.9 (Yang, 2007). To test whether the d_N/d_S ratio for each ortholog significantly deviated from 1, we conducted likelihood ratio tests of the likelihood estimates for the M1 model and the M0 model in which the d_N/d_S ratios were fixed to one, using a Δ ($\Delta = 2(\ln(M1) - \ln(M0))$) that approximates the chi-square distribution with one degree of freedom (Brand et al., 2015).

2.7. Antennal expression levels of orthologs and qRT-PCR verification

Comparative gene expression studies within a single species at different developmental stages or under different treatments have been widely reported whereas interspecific comparisons are rare and confined to a few closely related model species (Logacheva et al., 2011; Shiao et al., 2015). Here, given the close evolutionary relationship between *P. aenescens* and *P. maculicollis* and the similarity in transcriptome characteristics (e.g., unigene number, N50

length and mean sequence length; see Results), we made direct comparisons between the species to examine differences in the antennal expression levels of chemosensory genes. Based on the *de novo* transcriptomic data, we quantified the expression levels of chemosensory genes using the BGISEQ500 platform (BGI, Wuhan, China, <http://www.seq500.com/en/>) with 4 *P. aenescens* and 3 *P. maculicollis* biological replicates. The antennal tissue was collected, pooled, and treated using the same methods as described in section 2.1. To rule out the uncertainty in assignment of isoforms to orthologs, the expression profiles were only compared between the pairwise orthologs confirmed by the above ML analyses. Gene expression levels were estimated by RSEM (Li and Dewey, 2011). Specifically, we mapped the clean fragments back onto the assembled antennal libraries and obtained the fragment-count for each gene from the mapping results. The normalized value of fragments per kilobase of transcript per million mapped reads (FPKM) was used as a parameter to make comparisons of antennal expression levels between *P. aenescens* and *P. maculicollis* (Trapnell et al., 2010). We then identified the orthologs with significantly different expression using DESeq2 algorithm (Love et al., 2014), retaining only the orthologs with an absolute value of \log_2 fold-change >1 and a P_{adj} -value <0.05 (Benjamini and Hochberg, 1995). The expression profiles obtained from the RNA-seq were validated by qRT-PCR (Fig. S1). Specifically, we randomly selected twelve orthologs (OBP4, OBP10, OBP22, CSP2, CSP4, SNMP1a, OR11, OR15, OR16, GR4, GR5, IR2) on which to perform a quantitative real-time PCR (qRT-PCR) analysis using specific primers (Table S2). Two μ g of total RNA from each adult antennal sample was used to synthesize cDNA with the FastQuant RT Kit with gDNase (Tiangen, China) following the manufacturer's instructions. β -actin was used as internal references to normalize cDNA templates (Zhu et al., 2011). The qRT-PCR was performed on a PikoRea system (Thermo, USA) using SuperReal PreMix Plus (Tiangen, China), according to the manufacturer's instructions. The thermal cycling conditions were: 40 cycles at 95 °C for 10 s, 55 °C for 20 s, 72 °C for 30 s. The data were exported to EXCEL for a $2^{-\Delta\Delta CT}$ analysis with *P. maculicollis* as reference. Three independent biological replicates were performed for each treatment, and three technical replicates were performed for each reaction. We performed student t-tests to examine the difference in expression between the ratio of FPKM values of *P. aenescens* to *P. maculicollis* and the $2^{-\Delta\Delta CT}$ value using SPSS Statistics 19.0 (SPSS, Inc., Chicago, IL, USA).

3. Results

3.1. Transcriptome assembly

We combined the three samples (i.e., heads of newly emergent adults, heads of sexually mature adults, and antennae of sexually mature adults) to generate the assemblies for each species. Using this approach, we obtained a total of 59,642 unigenes with a total length of 49,050,523 bp, a mean length of 822 bp, and a N50 length of 1764 bp for *P. aenescens*. The *P. aenescens* assembly indicated that 20.62% of the unigenes (12,303 unigenes) were longer than 1000 bp. The results were similar for *P. maculicollis* where the assembly resulted in a total of 63,136 unigenes with a total length of 51,568,011 bp, a mean length of 817 bp, and a N50 length of 1679 bp. Approximately 21.5% of the unigenes (13,571 unigenes) were longer than 1000 bp (Table 1).

3.2. Annotation and functional classification

Searches of the Nr and Pfam databases identified a total of 23,329 and 26,277 annotated unigenes for *P. aenescens* and *P. maculicollis*, respectively. Of these, the number of unigenes over

Table 1

The assembly results of the transcriptomes of the two sister elm leaf beetles, *Pyrrhalta aenescens* and *P. maculicollis*. The number of unigenes in different length intervals are provided along with the percentage of all identified unigenes (in parentheses).

Length intervals	<i>Pyrrhalta aenescens</i>	<i>P. maculicollis</i>
200–300 bp	23,794 (39.89%)	24,867 (39.39%)
300–500 bp	14,454 (24.23%)	14,975 (23.72%)
500–1000 bp	9091 (15.24%)	9723 (15.40%)
>1000 bp	12,303 (20.62%)	13,571 (21.50%)
Total number of unigenes	59,642	63,136
Total length of unigenes	49,050,523	51,568,011
N50 length of unigenes	1764	1697
Mean length of unigenes	822.42	816.78

1000 bp in length was greater than 10,000 for both species (10,074 from *P. aenescens* and 11,314 from *P. maculicollis*; Table S3), and nearly 50% of the annotated unigenes were homologous to the coleopterans, *Tribolium castaneum* and *Dendroctonus ponderosae* (Fig. S2).

The functional classification analysis of gene ontology indicated that the transcriptomes of the sister beetle species are highly similar. For example, the ratio of annotated genes assigned to cellular components, molecular function and biological processes were 30.55%, 26.17% and 43.28% in *P. aenescens*, and 30.15%, 25.76% and 44.09% in *P. maculicollis* (Fig. 1). In addition, some unigenes that were highly abundant in both beetle species were classified into functions that might be involved in chemosensory perception in insects such as binding, transporter activity, receptor activity, and enzyme regulator activity (Fig. 1).

3.3. Chemosensory gene families

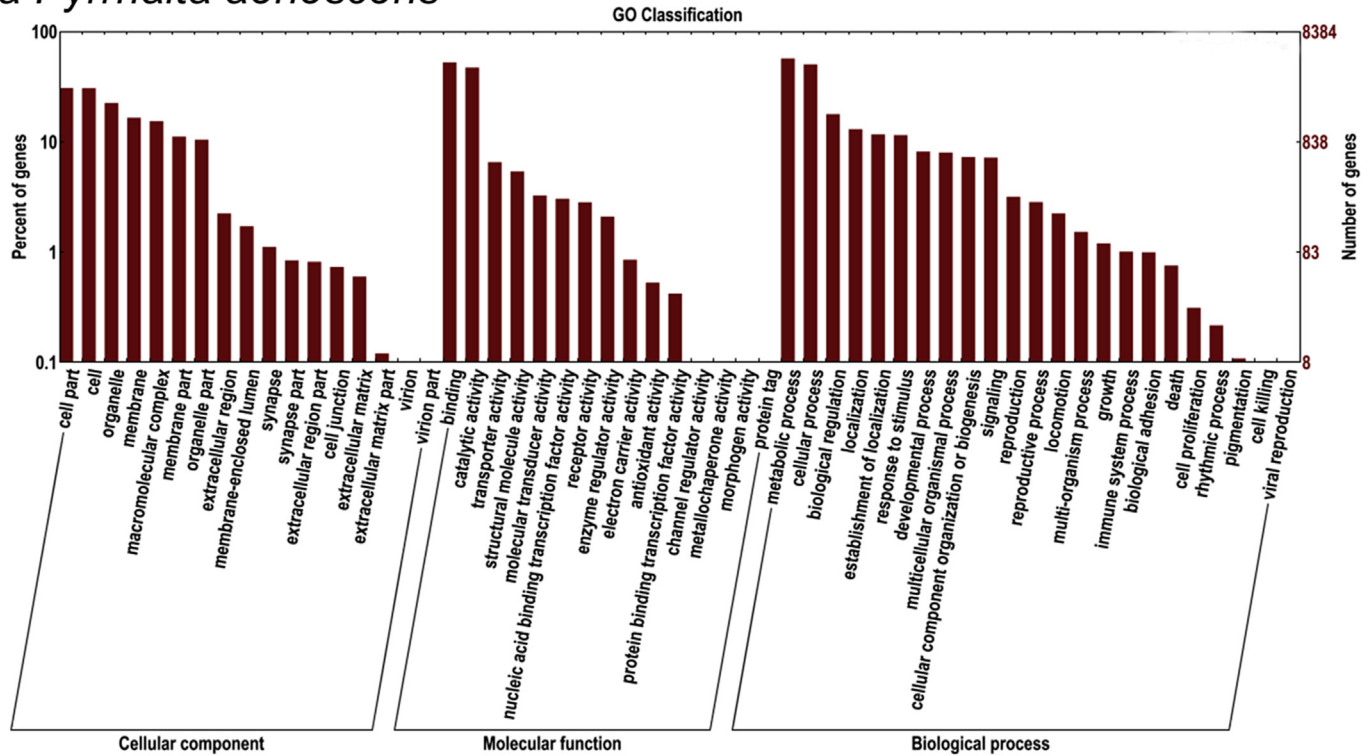
The annotated candidate chemosensory genes were filtered to remove the subset with Blastp results inconsistent with the annotation and candidates with E-values greater than 10^{-5} . In total, there were 31 OBPs, nine CSPs, two SNMPs, 26 ORs, 16 GRs, and eight IRs identified in *P. aenescens*. In *P. maculicollis*, we identified 36 OBPs, ten CSPs, two SNMPs, 22 ORs, ten GRs and seven IRs. For both species, the nucleotide or protein sequences in most of the putative chemosensory genes were complete or nearly full-length (Table 2). The identified olfactory genes were submitted to GenBank under the accession numbers KX298746–KX298837 and KX290605–KX290691.

3.3.1. Odorant binding proteins

Using their sequence properties, we classified the odorant binding proteins into three groups: Classic OBPs characterized by the presence of six cysteine residues at conserved positions, Plus-C OBPs with several additional cysteines and a diagnostic proline, and Minus-C OBPs with only four cysteine residues (Hekmat-Scafe et al., 2002; Sánchez-Gracia et al., 2009; Zhou et al., 2004). We identified eight Classic and 23 Minus-C OBPs in *P. aenescens*, and nine Classic and 27 Minus-C OBPs in *P. maculicollis*. Plus-C OBP, however, was not found in both of the two species (Fig. S3). The ML tree of OBPs from *Pyrrhalta* and published sequences from other model insects showed that the three groups of OBPs from *Pyrrhalta* beetles were dispersed throughout the tree with nearly half of them (18 in *P. aenescens* and 25 in *P. maculicollis*) located in the Minus-C lineage (Fig. 2).

The ML tree suggested 25 pairs of orthologous OBPs between the two beetle species, as well as six that were specific to *P. aenescens* and 11 to *P. maculicollis* (Fig. 2). The analysis of differential expression in these orthologs revealed that three orthologous pairs, OBP1 (\log_2 FC = -4.4144, $P < 0.001$), OBP12

a *Pyrrhalta aenescens*



b *Pyrrhalta maculicollis*

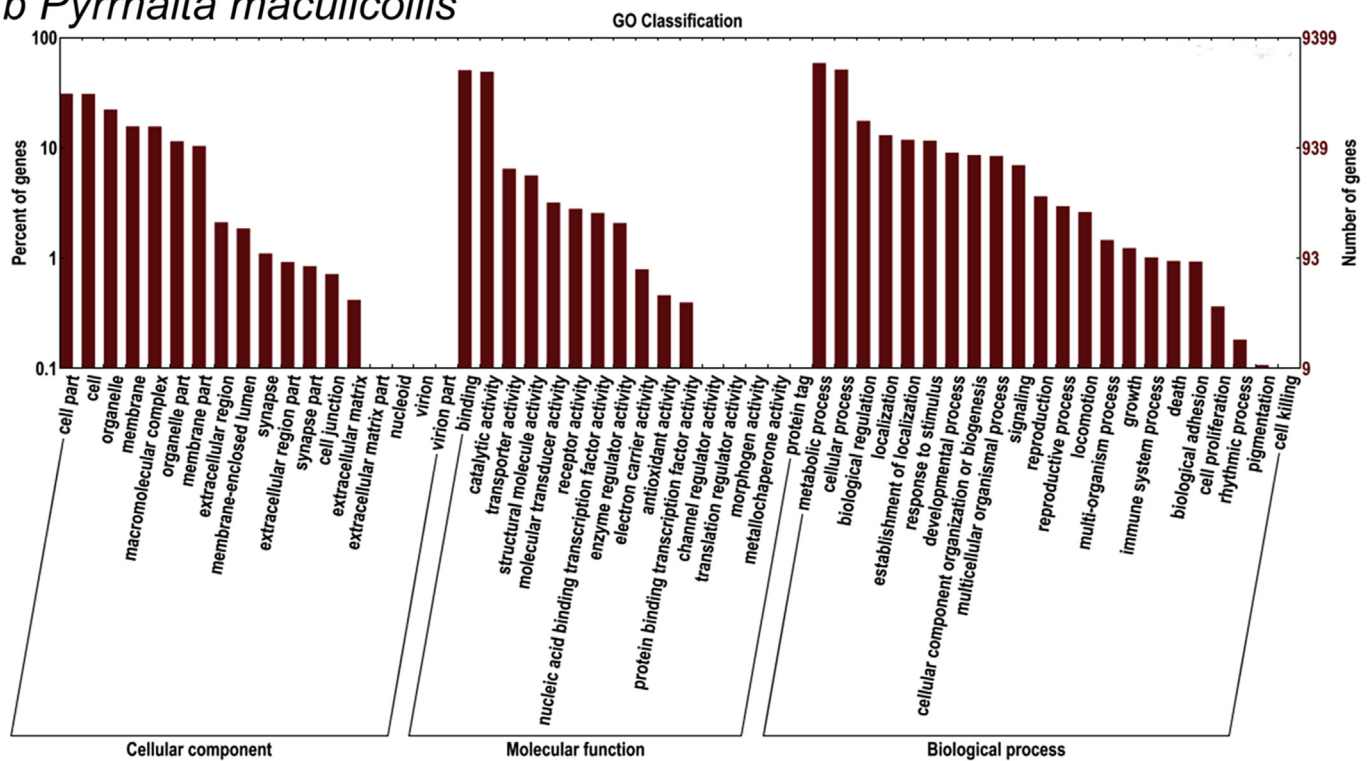


Fig. 1. Functional classification of the transcriptomes of the two sister beetle species, *Pyrrhalta aenescens* (a) and *P. maculicollis* (b). 8384 unigenes in *Pyrrhalta aenescens* and 9399 unigenes in *P. maculicollis* were examined through gene ontology (GO) and predicted to be involved in cellular components, molecular functions and biological processes.

($\log_2FC = 2.6033$, $P < 0.001$) and OBP25 ($\log_2FC = 2.8873$, $P = 0.0026$), had significantly different expression between two beetle species (Table 3).

The analysis of selection performed on the orthologs between the two sibling beetles indicated that only one ortholog, OBP10 ($d_N/d_S = 1.6022$, likelihood-ratio tests: $P = 0.0397$), was undergoing

Table 2

The number of chemosensory genes in the transcriptomes of the two sister elm leaf beetles, *Pyrrhalta aenescens* and *P. maculicollis*. The numbers in parentheses indicate the number of full-length, nearly full-length and partial length genes, respectively.

Species	Chemosensory gene families					
	OBPs	CSPs	SNMPs	ORs	GRs	IRs
<i>P. aenescens</i>	31 (20/11/0)	9 (7/2/0)	2 (1/1/0)	26 (11/14/1)	16 (9/4/3)	8 (3/5/0)
<i>P. maculicollis</i>	36 (22/14/0)	10 (7/3/0)	2 (1/1/0)	22 (12/8/2)	10 (5/3/2)	7 (3/2/2)

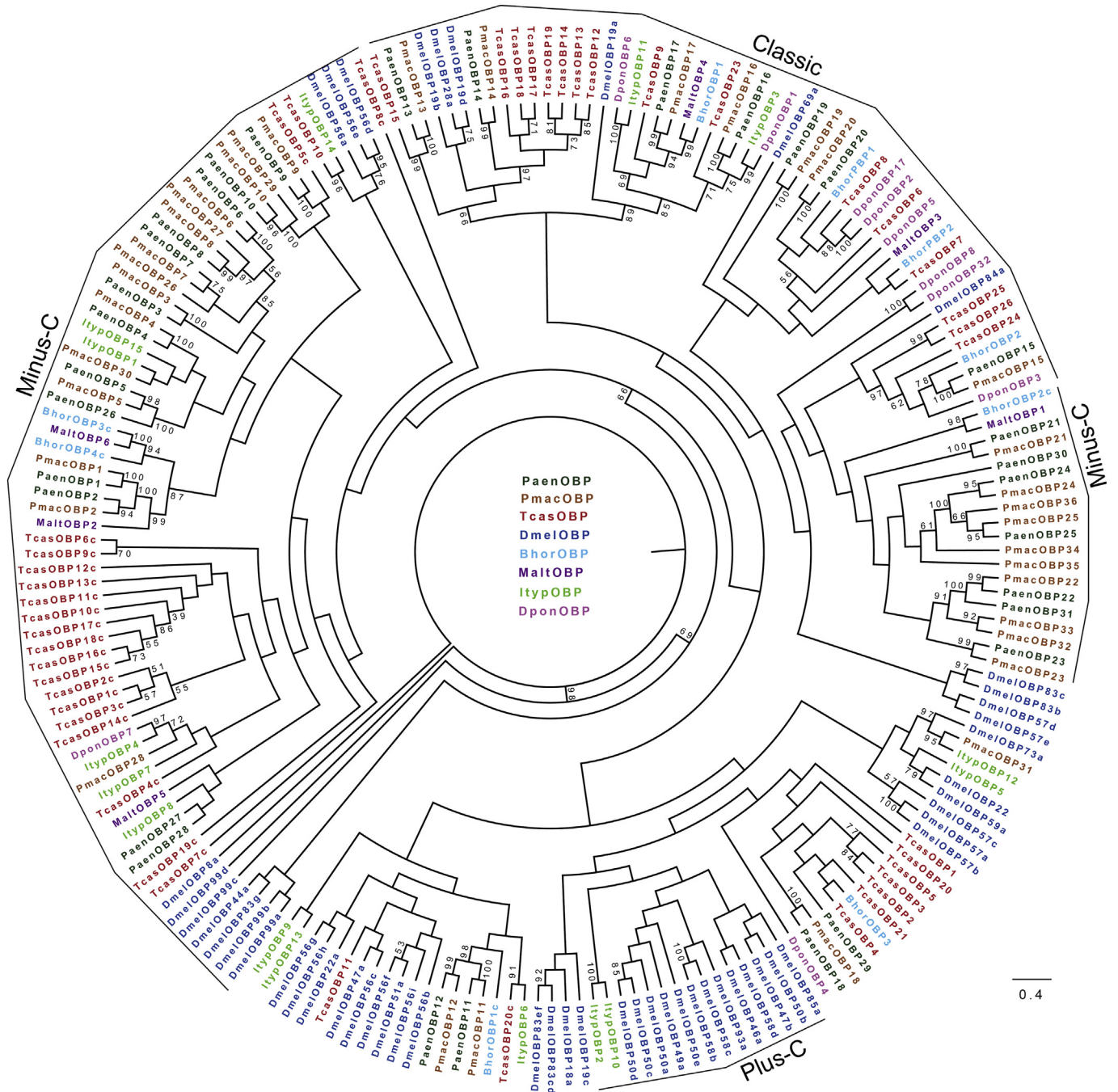


Fig. 2. Maximum-likelihood phylogeny based on protein sequences of candidate odorant binding proteins (OBPs). Included are OBPs from *Pyrrhalta aenescens* (Paen), *Pyrrhalta maculicollis* (Pmac), *Tribolium castaneum* (Tcas), *Drosophila melanogaster* (Dmel), *Ips typographus* (Ityp), *Batocera horsfieldi* (Bhor), *Monochamus alternatus* (Malt) and *Dendroctonus ponderosae* (Dpon). One major beetle-specific expansion of Minus-C OBPs is evident. Numbers refer to non-parametric bootstrap support (%).

Table 3

Candidate orthologs for chemosensory speciation between the two sister elm leaf beetles, *Pyrrhalta aenescens* and *P. maculicollis*, which either have significant difference in expression between each other or are under positive selection.

Orthologs	Different expression ^a		Selection pressure ^b		
	log ₂ FC	P _{adj}	d _N	d _S	d _N /d _S ^c
OBP1	-4.4144	5.36E-48	0.0056	0.0501	0.1108
OBP10	-0.3831	0.3772	0.0572	0.0357	1.6022*
OBP12	2.6033	6.31E-14	0.0134	0.0774	0.1731
OBP25	2.8873	0.0026	0.0584	0.2270	0.2573
CSP1	-3.7225	1.38E-17	0.0076	0.0699	0.1087
OR12	-1.9408	6.32E-10	0.0067	0.0525	0.1276
OR14	-1.7672	0.0101	0.0172	0.0351	0.4900
OR15	-3.3870	7.58E-35	0.0331	0.0231	1.4329*
GR4	2.3611	7.77E-10	0.0212	0.0524	0.4046

^a Calculated by the DESeq2 algorithm, negative values indicate higher expression in *P. aenescens*. Orthologs with an absolute value of log₂FC > 1 and P_{adj} < 0.05 were considered significantly different.

^b d_N: nonsynonymous substitution rate; d_S: synonymous substitution rate.

^c Orthologs with d_N/d_S significantly higher than 1 are indicated by *.

positive selection (Table 3). The remaining orthologs were under purifying selection as shown by d_N/d_S values that ranged from 0.0531 to 0.6525 (Dataset S2).

3.3.2. Chemosensory proteins

The ML analysis of the chemosensory proteins showed that there are seven pairs of orthologs shared by *P. aenescens* and *P. maculicollis*. In addition, two CSPs were unique to *P. aenescens* and three to *P. maculicollis* (Fig. S4). There was no significant difference in expression between orthologs except CSP1 (log₂FC = -3.7225, P < 0.001) (Table 3). The selection pressure analyses showed that purifying selection acted on all the CSP orthologs (Dataset S2).

3.3.3. Sensory neuron membrane proteins

There were two sensory neuron membrane proteins in each beetle species, constituting two orthologs (Fig. S5). The ML analysis showed that among insects there are two main groups of SNMPs, SNMP1 and SNMP2, and that all of the SNMPs in *Pyrrhalta* belong to the SNMP1 group (Fig. S5). Expression levels of the two orthologs were not significantly different between the species. The analysis of selection showed that both SNMP orthologs are under purifying selection (Dataset S2).

3.3.4. Odorant receptors

According to the ORs classification system of coleopteran insects based on Engsontia et al. (2008) and Andersson et al. (2013), a ML phylogenetic tree was constructed with ORs from *Pyrrhalta* and representative ORs from other beetle species (Fig. 3). We identified 16 pairs of orthologous odorant receptors in the two sister beetles, ten unique to *P. aenescens* and six unique to *P. maculicollis* (Fig. 3). Consistent with the two previous studies, seven groups were recovered in our trees. Specifically, 1 and 2 are the common groups containing a mixture of ORs from different species; groups 3, 4, 5, and 6 are mainly *Tribolium castaneum*-specific except for a few group 3 ORs that were also found in *Pyrrhalta* (four from *P. aenescens* and five from *P. maculicollis*). Group 7 is a specific expansion in bark beetles (DponOR and ItypOR). As compared to other beetle species, the *Pyrrhalta* studied here have no specific group, but are mainly scattered throughout common group 2 (14 in *P. aenescens* and 10 in *P. maculicollis*). The conserved olfactory co-receptor *Orcos* were clustered; however, none of these receptors were identified in *Pyrrhalta* (Fig. 3).

The expression analysis showed that three orthologs, OR12 (log₂FC = -1.9408, P < 0.001), OR14 (log₂FC = -1.7672, P = 0.0101)

and OR15 (log₂FC = -3.3870, P < 0.001), were differentially expressed between the two sister leaf beetles (Table 3).

The analysis of OR orthologs between *P. aenescens* and *P. maculicollis* showed that most of these genes are under purifying selection (d_N/d_S: 0.1485–0.8514) (Dataset S2). Only OR15 was under positive selection (d_N/d_S = 1.4329, likelihood-ratio tests: P = 0.0207) (Table 3).

3.3.5. Gustatory receptors

ML analysis suggested that there are eight pairs of orthologous gustatory receptors between the species, eight that are unique to *P. aenescens*, and two that are unique to *P. maculicollis* (Fig. S6). Only one ortholog, GR4 (log₂FC = 2.3611, P < 0.001), was differentially expressed between *P. aenescens* and *P. maculicollis* (Table 3). The ML tree also identified a *Pyrrhalta* specific clade and two putative CO₂ receptors (PaenGR3 and PaenGR12) that clustered with those from other insects (Fig. S6). The d_N/d_S values of these orthologs ranged from 0.1046 to 0.5921, showing purifying selection (Dataset S2).

3.3.6. Ionotropic receptors

We found four pairs of orthologous ionotropic receptors shared by the beetle species, as well as four unique to *P. aenescens*, and three to *P. maculicollis* (Fig. S7). We observed no difference in gene expression between the species (Table 3). According to their conservatism and function, IRs can be divided into two groups: antennal IRs, which are conserved across insect orders and functioned in olfaction, and divergent IRs, which are usually a species-specific expansion and might have a role in taste (Croset et al., 2010). The ML tree showed that most of the IR genes in these two sister species (PaenIR1, PaenIR2, PaenIR4, PaenIR5, PaenIR6 in *P. aenescens* and PmaclR1, PmaclR2, PmaclR4, in *P. maculicollis*) are putative conserved antennal IRs, which were clustered with IR21a, IR64a and IR75a of *Tribolium castaneum* and *Drosophila melanogaster* (Fig. S7). All of these orthologs were found to be evolving under purifying selection (Dataset S2).

4. Discussion

Divergence in chemosensory traits is likely a key component of speciation in many insects as changes in these traits can influence insect behaviors that subsequently mediate reproductive isolation. Although identification of changes in chemosensory gene expression is an important starting point for understanding how chemosensory traits contribute to diversification, there are surprisingly few studies that compare sister taxa differing in host plant use. Moreover, taxonomic sampling within groups also remains quite narrow. For example, among Coleoptera, only four species have been characterized with respect to the repertoires of OBPs, CSPs, SNMPs, ORs, GRs and IRs: the flour beetle *T. castaneum* (Engsontia et al., 2008), a scarab beetle *Hylamorphia elegans* (Venthur et al., 2016), and two bark beetles, *I. typographus* and *D. ponderosae* (Andersson et al., 2013). Here we use high throughput transcriptomic analyses to identify the chemosensory gene families in a sister species pair of leaf beetles in the genus *Pyrrhalta*, providing the first view on chemosensory genes in the Chrysomelidae. Despite the overall low divergence in transcriptome profiles between these sibling species, the results suggest positive selection on chemosensory traits as there was evidence of divergent evolution of some chemosensory genes. These divergent genes, therefore, might contribute at least partially to chemical-mediated speciation between these sister species.

4.1. Transcriptome overview

As was the expectation for comparisons of closely related insect

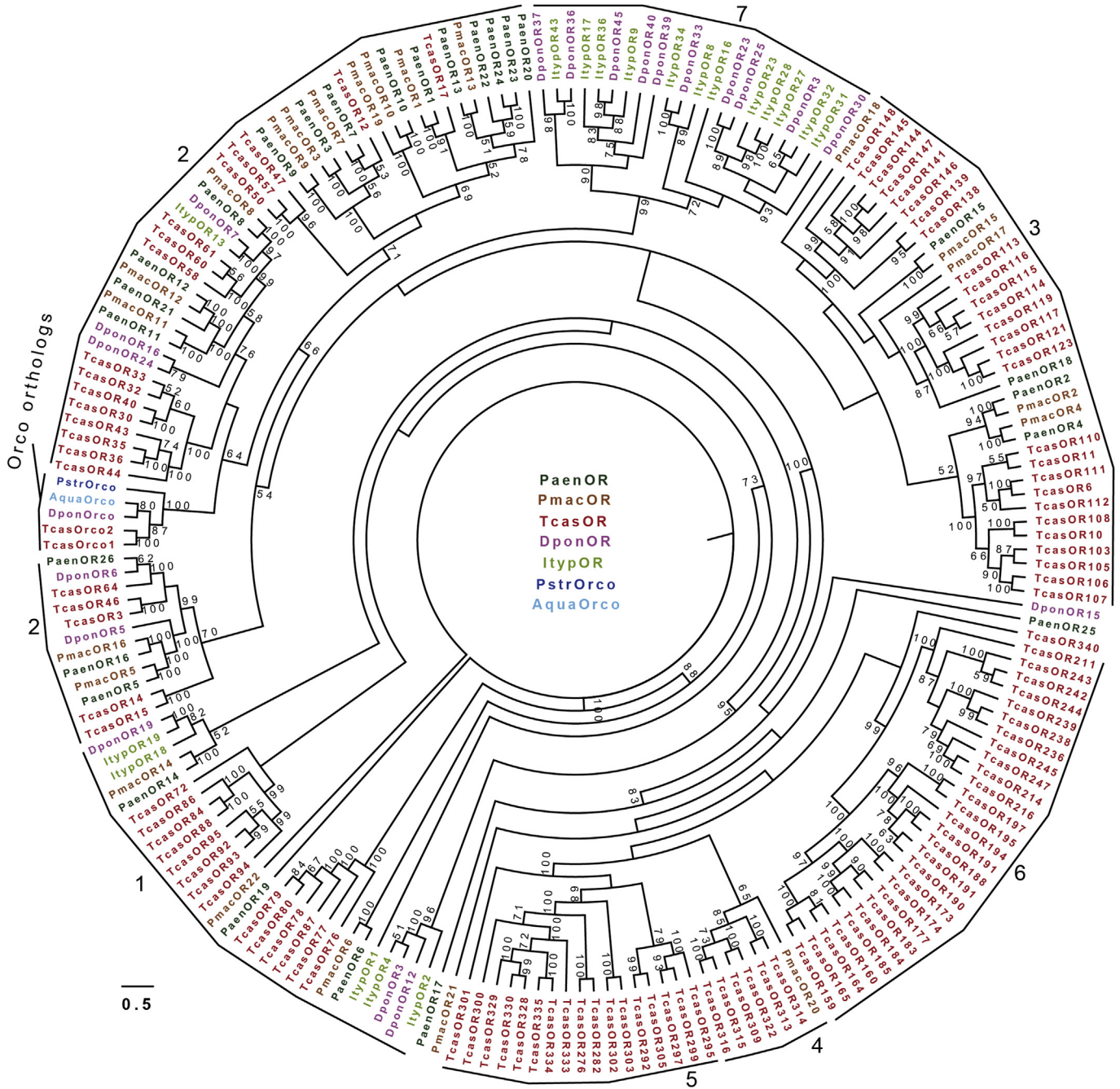


Fig. 3. Maximum-likelihood phylogeny based on protein sequences of candidate odorant receptors (ORs). Included are ORs from *Pyrrhalta aenescens* (Paen), *Pyrrhalta maculicollis* (Pmac), *Tribolium castaneum* (Tcas), *Drosophila melanogaster* (Dmel), *Ips typographus* (Ityp), *Dendroctonus ponderosae* (Dpon), *Phyllotreta striolata* (Pstr) and *Ambrostoma quadriimpressum* (Aqua). Seven lineages were identified according to the OR classification system of Coleopteran insects based on Engson et al. (2008) and Andersson et al. (2013). Numbers refer to non-parametric bootstrap support (%).

species, we observed high similarity in expression patterns for *P. aenescens* and *P. maculicollis*. The transcriptome assembly revealed consistent results between the species in terms of assembly parameters such as the number of unigenes, N50 length, and mean length. In addition, the number of annotated genes, GO functional classification, and the species distribution of hit orthologs were also highly similar. These results uphold patterns previously observed in other sibling species. For example, transcriptomic analysis showed a similar pattern in gene number and annotation results between *Ostrinia nubilalis* and its sister species *O. scapularis* (Gschloessl et al., 2013), and also between *Dendrolimus houi* and

D. kikuchii (Zhang et al., 2014b). Moreover, the overall patterns observed in the present transcriptome assembly corroborate previous molecular phylogenetic findings indicating that *P. aenescens* and *P. maculicollis* are sister species (Nie et al., 2012).

4.2. Chemosensory gene families

4.2.1. Non-receptor gene families

The first critical step of chemosensation in insects is mediated by the OBPs and CSPs; thus, these gene families have received considerable attention (Danty et al., 1998; Gong et al., 2009;

Graham and Davies, 2002; Jin et al., 2005; Pelosi et al., 2006; Vogt et al., 1991a,b; Zhou et al., 2009). When chemical signals enter into the sensillum lymph, OBPs and/or CSPs act as carriers that transport the signals to the receptors. Our survey for these gene families in *Pyrrhalta* revealed a total of 31 and 36 OBPs and 9 and 10 CSPs in *P. aenescens* and *P. maculicollis*, respectively. These totals are similar to those observed in the *Dendroctonus ponderosae* transcriptome (OBPs: 31, CSPs: 11) (Andersson et al., 2013), but are far fewer than those found in *Tribolium castaneum* (OBPs: 49, CSPs: 20) (Engsontia et al., 2008). One caveat to consider is that the number of genes detected will vary between studies as a consequence of differences in RNA-seq technology, analysis methods, and sampling. For instance, some genes could have been missed in our analysis if they were expressed in non-olfactory or gustatory tissues (Pelosi et al., 2006; Wanner et al., 2005) or at different life history stages (Engsontia et al., 2008; Xia et al., 2008), especially given that our sampling design used only the heads or antennae of adult beetles. This point aside, the number of genes identified in our analysis seems reasonable given previous estimates in other beetle species.

Similar to other work in beetles, we found an expansion of the Minus-C OBPs in *Pyrrhalta*, a pattern that has been identified in *Tribolium* and bark beetles (Andersson et al., 2013). Analysis of the properties of the OBP amino acid sequences revealed 23 and 27 Minus-C OBPs in *P. aenescens* and *P. maculicollis*, respectively. An expansion of the Minus-C OBPs suggests that these genes might play an important role in chemosensory functions in Coleoptera.

The final non-receptor gene family class that we examined is the SNMP, a group that contains two sub-families (SNMP1 and SNMP2) (Robertson et al., 1999; Rogers et al., 1997). We identified two SNMP1s in each species and both species entirely lacked members of the SNMP2 sub-family. This pattern may be explained by the functional importance of SNMP1s, that are typically expressed in olfactory receptor neurons and are involved in pheromone recognition (Benton et al., 2007; Vogt et al., 2009). The absence of SNMP2s requires further confirmation because it is possible that this sub-family may be expressed during other developmental stages or in different tissues.

4.2.2. Receptor gene families

Although ORs play a key role in the olfaction of insects, we found fewer ORs than previous work on other insects. For example, we identified 26 ORs in *P. aenescens* and 22 in *P. maculicollis*, versus the numbers identified in other coleopterans such as *I. typographus* (43) and *D. ponderosae* (49) bark beetles (Andersson et al., 2013), the longhorned beetle *Megacyllene caryae* (57) (Mitchell et al., 2012), and the flour beetle *T. castaneum* (341) (Engsontia et al., 2008). This pattern suggests an expansion in *Tribolium* as compared to other beetles. Additional work is needed to appreciate whether these patterns are upheld.

In contrast to the ORs, we identified more GRs than previous studies of beetles. In the present study, we identified 16 (*P. aenescens*) and 10 (*P. maculicollis*) GRs as compared to those found in *I. typographus* (2) and *D. ponderosae* (6) (Andersson et al., 2013). This pattern likely occurred because our samples comprised both antennae and heads, whereas the bark beetle transcriptomes were derived from antennae only (Andersson et al., 2013). In addition, we found two putative CO₂ receptors (PaenGR3 and PaenGR12) in *P. aenescens*, suggesting that additional CO₂ receptors exist in this species because this protein only functions as a heterodimer or heterotrimer for proper CO₂ recognition (Jones et al., 2007; Lu et al., 2007; Robertson and Kent, 2009). Whether these receptors exist in *P. maculicollis* remains to be determined.

Finally, among the IRs identified in *P. aenescens* (8) and *P. maculicollis* (7), five in *P. aenescens* and three in *P. maculicollis* were conserved, putative antennal IRs, homologous to IR21a, IR64a

and IR75a in *T. castaneum* and *D. melanogaster* (Croset et al., 2010). In contrast to the expansion of the divergent IRs in other insects (Croset et al., 2010), we observed few divergent IRs in *Pyrrhalta*.

4.3. Candidate genes for chemosensory speciation

Divergent chemosensory genes linked to differences in the sensory tuning of sister species represent a response to divergent selection on chemosensory traits (Smadja and Butlin, 2009); thus, these genes are candidates for chemical-mediated speciation. For instance, comparison of expression of chemosensory gene families between the sister beetle species suggested several candidate genes that may have been involved in the speciation process. A number of genes were identified that were expressed in only one of the species or that had strong differences in expression level between *P. aenescens* and *P. maculicollis* (e.g., OBP1, OBP12, OBP25, CSP1, OR12, OR14, OR15 and GR4). Furthermore, there were also changes in gene sequence that suggested two genes were undergoing positive selection (OBP10 and OR15). Interestingly, the OR15 ortholog also had significant differential expression between *P. aenescens* and *P. maculicollis*. This pattern was particularly exciting because it suggests that these genes might play a key role in physiological or ecological functions in *Pyrrhalta* that might contribute to reproductive isolation.

Myriad ecological factors may have promoted differentiation of chemosensory genes between sympatric sister species, including changes in mating ecology (Leary et al., 2012) and host shifts (Smadja et al., 2009, 2012). For the sister beetle species studied here, previous studies have demonstrated that the difference in female cuticular hydrocarbons have caused divergence in male mate recognition (Zhang et al., 2014a), and that differences in leaf wax profiles of plants differing in age could also lead to divergent selection (Zhang et al., 2015). Thus, it is likely that the divergence in expression of chemosensory genes observed in the present study is due to selection acting on both host specialization and mating behaviors, leading to habitat isolation and sexual isolation between sympatric sibling species.

5. Conclusions

Here we performed a comprehensive analysis of antennal transcriptomes between *P. aenescens* and *P. maculicollis*, focused on chemosensory genes. Despite the overall low divergence in transcriptomes, several candidate genes were identified that might contribute to reproductive isolation between these sister species. Particularly for highly and differently expressed genes under positive selection, we predict that these chemosensory genes play a key role in chemical-based sexual and habitat isolation in *Pyrrhalta*. Combined with the available chemical and ecological work in this system, further studies of these divergent chemosensory genes presented here will provide insight into the process of chemosensory speciation among *Pyrrhalta* beetles.

Authors' contributions

BZ, XKY and HJX conceived and designed the experimental plan. BZ, WZ and HJX carried out the laboratory work and data analysis. BZ, HJX, REN and WZL collected the insects in the field. BZ, KAS and HJX drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This study was supported by National Natural Science Foundation of China (Grant No. 31472030).

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.ibmb.2016.11.001>.

References

- Abuin, L., Bargeton, B., Ulbrich, M.H., Isacoff, E.Y., Kellenberger, S., Benton, R., 2011. Functional architecture of olfactory ionotropic glutamate receptors. *Neuron* 69, 44–60.
- Andersson, M.N., Grosse-Wilde, E., Keeling, C.I., Bengtsson, J.M., Yuen, M.M., Li, M., Hillbur, Y., Bohlmann, J., Hansson, B.S., Schlyter, F., 2013. Antennal transcriptome analysis of the chemosensory gene families in the tree killing bark beetles, *Ips typographus* and *Dendroctonus ponderosae* (Coleoptera: Curculionidae: Scolytinae). *BMC Genomics* 14, 198.
- Bengtsson, J.M., Trona, F., Montagné, N., Anfora, G., Ignell, R., Witzgall, P., Jacquinjoly, E., 2012. Putative chemosensory receptors of the codling moth, *Cydia pomonella*, identified by antennal transcriptome analysis. *PLoS One* 7, e31620.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300.
- Benton, R., Vannice, K.S., Vossahl, L.B., 2007. An essential role for a CD36-related receptor in pheromone detection in *Drosophila*. *Nature* 450, 289–293.
- Benton, R., Vannice, K.S., Gomez-Diaz, C., Vossahl, L.B., 2009. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* 136, 149–162.
- Boake, C.R.B., 1991. Coevolution of senders and receivers of sexual signals: genetic coupling and genetic correlations. *Trends Ecol. Evol.* 6, 225–227.
- Brand, P., Ramirez, S.R., Leese, F., Quezada-Euan, J.J., Tollrian, R., Eltz, T., 2015. Rapid evolution of chemosensory receptor genes in a pair of sibling species of orchid bees (Apidae: Euglossini). *BMC Evol. Biol.* 15, 176.
- Caillaud, M.C., Via, S., 2000. Specialized feeding behavior influences both ecological specialization and assortative mating in sympatric host races of pea aphids. *Am. Nat.* 156, 606–621.
- Carey, A.F., Wang, G., Su, C.-Y., Zwiebel, L.J., Carlson, J.R., 2010. Odorant reception in the malaria mosquito *Anopheles gambiae*. *Nature* 464, 66–71.
- Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676.
- Croset, V., Rytz, R., Cummins, S.F., Budd, A., Brawand, D., Kaessmann, H., Gibson, T.J., Benton, R., 2010. Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Gene* 6, e1001064.
- Danty, E., Arnold, G., Huet, J.C., Huet, D., Masson, C., Pernellet, J.C., 1998. Separation, characterization and sexual heterogeneity of multiple putative odorant-binding proteins in the honeybee *Apis mellifera* L. (Hymenoptera: Apidae). *Chem. Sens.* 23, 83–91.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2011. ProfTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 27, 1164–1165.
- Dobritsa, A.A., van der Goes van Naters, W., Warr, C.G., Steinbrecht, R.A., Carls, J.R., 2003. Integrating the molecular and cellular basis of odor coding in the *Drosophila* antenna. *Neuron* 37, 827–841.
- Ehrlich, P.R., Raven, P.H., 1964. Butterflies and plants: a study in coevolution. *Evolution* 18, 586–608.
- Engstontia, P., Sanderson, A.P., Cobb, M., Walden, K.K., Robertson, H.M., Brown, S., 2008. The red flour beetle's large nose: an expanded odorant receptor gene family in *Tribolium castaneum*. *Insect Biochem. Mol. Biol.* 38, 387–397.
- Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L.L., Tate, J., Punta, M., 2014. Pfam: the protein families database. *Nucleic Acids Res.* 42, d222–d230.
- Funk, D.J., 1998. Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. *Evolution* 52, 1744–1759.
- Gong, D.P., Zhang, H.J., Zhao, P., Xia, Q.Y., Xiang, Z.H., 2009. The odorant binding protein gene family from the genome of silkworm, *Bombyx mori*. *BMC Genomics* 10, 332.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q.D., Chen, Z.H., Mucic, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652.
- Graham, L.A., Davies, P.L., 2002. The odorant-binding proteins of *Drosophila melanogaster*: annotation and characterization of a divergent gene family. *Gene* 292, 43–55.
- Grosse-Wilde, E., Kuebler, L.S., Bucks, S., Vogel, H., Wicher, D., Hansson, B.S., 2011. Antennal transcriptome of *Manduca sexta*. *Proc. Natl. Acad. Sci. U. S. A.* 108, 7449–7454.
- Gschloessl, B., Beyne, E., Audiot, P., Bourguet, D., Streiff, R., 2013. *De novo* transcriptomic resources for two sibling species of moths: *Ostrinia nubilalis* and *O. scapularis*. *BMC Res. Notes* 6, 73.
- Gu, S.H., Zhou, J.J., Gao, S., Wang, D.H., Li, X.C., Guo, Y.Y., Zhang, Y.J., 2015. Identification and comparative expression analysis of odorant binding protein genes in the tobacco cutworm *Spodoptera litura*. *Sci. Rep.* 5.
- Hallem, E.A., Carlson, J.R., 2006. Coding of odors by a receptor repertoire. *Cell* 125, 143–160.
- Hekmat-Scafe, D.S., Scafe, C.R., McKinney, A.J., Tanouye, M.A., 2002. Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. *Genome Res.* 12, 1357–1369.
- Ishii, K., Hirai, Y., Katagiri, C., Kimura, M.T., 2001. Sexual isolation and cuticular hydrocarbons in *Drosophila elegans*. *Heredity* 87, 392–399.
- Janz, N., Nylin, S., Wahlberg, N., 2006. Diversity begets diversity: host expansions and the diversification of plant-feeding insects. *BMC Evol. Biol.* 6, 4.
- Jin, X., Brandazza, A., Navarrini, A., Ban, L., Zhang, S., Steinbrecht, R.A., Zhang, L., Pelosi, P., 2005. Expression and immunolocalisation of odorant binding and chemosensory proteins in locusts. *Cell. Mol. Life Sci.* 62, 1156–1166.
- Jones, W.D., Cayirlioglu, P., Kadow, I.G., Vossahl, L.B., 2007. Two chemosensory receptors together mediate carbon dioxide detection in *Drosophila*. *Nature* 445, 86–90.
- Katoh, K., Misawa, K., Kuma, K.-I., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Res.* 30, 3059–3066.
- Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- Kaupp, U.B., 2010. Olfactory signalling in vertebrates and insects: differences and commonalities. *Nat. Rev. Neurosci.* 11, 188–200.
- Kitabayashi, A.N., Arai, T., Kubo, T., Natori, S., 1998. Molecular cloning of cDNA for p10, a novel protein that increases in the regenerating legs of *Periplaneta americana* (American cockroach). *Insect Biochem. Mol. Biol.* 28, 785–790.
- Kwon, J.Y., Dahanukar, A., Weiss, L.A., Carlson, J.R., 2007. The molecular basis of CO₂ reception in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 3574–3578.
- Leal, W.S., 2013. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* 58, 373–391.
- Leary, G.P., Allen, J.E., Bunger, P.L., Luginbill, J.B., Linn, C.E., Macallister, I.E., Kavanaugh, M.P., Wanner, K.W., 2012. Single mutation to a sex pheromone receptor provides adaptive specificity between closely related moth species. *Proc. Natl. Acad. Sci. U. S. A.* 109, 14081–14086.
- Legeai, F., Malpel, S., Montagné, N., Monsemper, C., Cusserans, F., Merlin, C., François, M.C., Maibèche-Coisné, M., Gavory, F., Poulain, J., Jacquinjoly, E., 2011. An expressed sequence tag collection from the male antennae of the noctuid moth *Spodoptera littoralis*: a resource for olfactory and pheromone detection research. *BMC Genomics* 12, 86.
- Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* 12, 323.
- Löfstedt, C., Herrebout, W.M., Menken, S.B.J., 1991. Sex pheromones and their potential role in the evolution of reproductive isolation in small ermine moths (Yponomeutidae). *Chemoecology* 2, 20–28.
- Logacheva, M.D., Kasianov, A.S., Vinogradov, D.V., Samigullin, T.H., Gelfand, M.S., Makeev, V.J., Penin, A.A., 2011. *De novo* sequencing and characterization of floral transcriptome in two species of buckwheat (*Fagopyrum*). *BMC Genomics* 12, 30.
- Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA seq data with DESeq2. *Genome Biol.* 15, 550.
- Lu, T., Qiu, Y.T., Wang, G., Kwon, J.Y., Rutzler, M., Kwon, H.W., Pitts, R.J., van Loon, J.J.A., Takken, W., Carlson, J.R., Zwiebel, L.J., 2007. Odor coding in the maxillary palp of the malaria vector mosquito *Anopheles gambiae*. *Curr. Biol.* 17, 1533–1544.
- McBride, C.S., 2007. Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 4996–5001.
- Mitchell, R.F., Hughes, D.T., Luetje, C.W., Millar, J.G., Soriano-Agaton, F., Hanks, L.M., Robertson, H.M., 2012. Sequencing and characterizing odorant receptors of the cerambycid beetle *Megaclypea caryae*. *Insect Biochem. Mol. Biol.* 42, 499–505.
- Nie, R.E., Xue, H.J., Hua, Y., Yang, X.K., Vogler, A.P., 2012. Distinct species or colour polymorphism? Life history, morphology and sequence data separate two *Pyrrhalta* elm beetles (Coleoptera: Chrysomelidae). *Syst. Biodivers.* 10, 133–146.
- Nosil, P., Crespi, B.J., Gries, R., Gries, G., 2007. Natural selection and divergence in mate preference during speciation. *Genetica* 129, 309–327.
- Oppenheim, S.J., Baker, R.H., Simon, S., DeSalle, R., 2015. We can't all be supermodels: the value of comparative transcriptomics to the study of non-model insects. *Insect Mol. Biol.* 24, 139–154.
- Pelosi, P., Zhou, J.J., Ban, L.P., Calvello, M., 2006. Soluble proteins in insect chemical communication. *Cell. Mol. Life Sci.* 63, 1658–1676.
- Peterson, M.A., Dobler, S., Larson, E.L., Juárez, D., Schlarbaum, T., Monsen, K.J., Francke, W., 2007. Profiles of cuticular hydrocarbons mediate male mate choice and sexual isolation between hybridizing *Chrysochus* (Coleoptera: Chrysomelidae). *Chemoecology* 17, 87–96.
- Robertson, H.M., Martos, R., Sears, C.R., Todres, E.Z., Walden, K.K., Nardi, J.B., 1999. Diversity of odourant binding proteins revealed by an expressed sequence tag project on male *Manduca sexta* moth antennae. *Insect Mol. Biol.* 8, 501–518.
- Robertson, H.M., Kent, L.B., 2009. Evolution of the gene lineage encoding the carbon dioxide receptor in insects. *J. Insect Sci.* 9, 19.
- Rogers, M.E., Sun, M., Lerner, M.R., 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.
- Sakurai, T., Nakagawa, T., Mitsuno, H., Mori, H., Endo, Y., Tanoue, S., Yasukochi, Y., Touhara, K., Nishioka, T., 2004. Identification and functional characterization of a sex pheromone receptor in the silkworm *Bombyx mori*. *Proc. Natl. Acad. Sci. U.*

- S. A. 101, 16653–16658.
- Sánchez-Gracia, A., Vieira, F.G., Rozas, J., 2009. Molecular evolution of the major chemosensory gene families in insects. *Heredity* 103, 208–216.
- Shiao, M.S., Chang, J.M., Fan, W.L., Lu, M.Y., Notredame, C., Fang, S., Kondo, R., Li, W.H., 2015. Expression divergence of chemosensory genes between *Drosophila sechellia* and its sibling species and its implications for host shift. *Genome Biol. Evol.* 7, 2843–2858.
- Smadja, C., Butlin, R.K., 2009. On the scent of speciation: the chemosensory system and its role in premating isolation. *Heredity* 102, 77–97.
- Smadja, C., Shi, P., Butlin, R.K., Robertson, H.M., 2009. Large gene family expansions and adaptive evolution for odorant and gustatory receptors in the pea aphid, *Acyrtosiphon pisum*. *Mol. Biol. Evol.* 26, 2073–2086.
- Smadja, C.M., Canbäck, B., Vitalis, R., Gautier, M., Ferrari, J., Zhou, J.J., Butlin, R.K., 2012. Large-scale candidate gene scan reveals the role of chemoreceptor genes in host plant specialization and speciation in the pea aphid. *Evolution* 66, 2723–2738.
- Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.
- Stensmyr, M.C., Dweck, H.K.M., Farhan, A., Ibba, I., Strutz, A., Mukunda, L., Linz, J., Grabe, V., Steck, K., Lavista-Llanos, S., Wicher, D., Sachse, S., Knaden, M., Becher, P.G., Seki, Y., Hansson, B.S., 2012. A conserved dedicated olfactory circuit for detecting harmful microbes in *Drosophila*. *Cell* 151, 1345–1357.
- Stireman, J.O., Nason, J.D., Heard, S.B., 2005. Host-associated genetic differentiation in phytophagous insects: general phenomenon or isolated exceptions? Evidence from a goldenrod-insect community. *Evolution* 59, 2573–2587.
- Suyama, M., Torrents, D., Bork, P., 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34, W609–W612.
- Touhara, K., Vossell, L.B., 2009. Sensing odorants and pheromones with chemosensory receptors. *Annu. Rev. Physiol.* 71, 307–332.
- Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., Pachter, L., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515.
- Ventur, H., Zhou, J.J., Mutis, A., Ceballos, R., Mella-Herrera, R., Larama, G., Avila, A., Iturriaga-Vásquez, P., Faundez-Parraguez, M., Alvear, M., Quiroz, A., 2016. β -Ionone as putative semiochemical suggested by ligand binding on an odorant binding protein of *Hyalomorpha elegans* and electroantennographic recordings. *Entomol. Sci.* 19, 188–200.
- Vereecken, N.J., Mant, J., Schiestl, F.P., 2007. Population differentiation in female sex pheromone and male preferences in a solitary bee. *Behav. Ecol. Sociobiol.* 61, 811–821.
- Vogt, R.G., Prestwich, G.D., 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., 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., Miller, N.E., Litvack, R., Fandino, R.A., Sparks, J., Staples, J., Friedman, R., Dickens, J.C., 2009. The insect SNMP gene family. *Insect Biochem. Mol. Biol.* 39, 448–456.
- Vogt, R.G., Große-Wilde, E., Zhou, J.J., 2015. The Lepidoptera odorant binding protein gene family: gene gain and loss within the GOBP/PBP complex of moths and butterflies. *Insect Biochem. Mol. Biol.* 62, 142–153.
- Vosshall, L.B., Stocker, R.F., 2007. Molecular architecture of smell and taste in *Drosophila*. *Annu. Rev. Neurosci.* 30, 505–533.
- Wanner, K.W., Isman, M.B., Feng, Q., Plettner, E., Theilmann, D.A., 2005. Developmental expression patterns of four chemosensory protein genes from the Eastern spruce budworm, *Chroistoneura fumiferana*. *Insect Mol. Biol.* 14, 289–300.
- Xia, Y., Wang, G., Buscariollo, D., Pitts, R.J., Wenger, H., Zwiebel, L.J., 2008. The molecular and cellular basis of olfactory-driven behavior in *Anopheles gambiae* larvae. *Proc. Natl. Acad. Sci. U. S. A.* 105, 6433–6438.
- Yang, Z.H., 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591.
- Zhang, B., Xue, H.J., Song, K.Q., Nie, R.E., Li, W.Z., Yang, X.K., 2015. Adaptation to different host plant ages facilitates insect divergence without a host shift. *Proc. R. Soc. Lond. B* 282, 20151649.
- Zhang, B., Xue, H.J., Song, K.Q., Liu, J., Li, W.Z., Nie, R.E., Yang, X.K., 2014a. Male mate recognition via cuticular hydrocarbons facilitates sexual isolation between sympatric leaf beetle sister species. *J. Insect Physiol.* 70, 15–21.
- Zhang, C., Bailey, D.K., Awad, T., Liu, G.Y., Xing, G.L., Cao, M.Q., Valmeekam, V., Retief, J., Matsuzaki, H., Taub, M., Seielstad, M., Kennedy, G.C., 2006. A whole genome long-range haplotype (WGLRH) test for detecting imprints of positive selection in human populations. *Bioinformatics* 22, 2122–2128.
- Zhang, S.F., Zhang, Z., Wang, H.B., Kong, X.B., 2014b. Antennal transcriptome analysis and comparison of olfactory genes in two sympatric defoliators, *Dendrolimus houi* and *Dendrolimus kikuchii* (Lepidoptera: Lasiocampidae). *Insect Biochem. Mol. Biol.* 52, 69–81.
- Zhou, J.J., Huang, W., Zhang, G.A., Pickett, J.A., 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., Robertson, G., He, X., Dufour, S., Hooper, A.M., Pickett, J.A., Keep, N.H., Field, L.M., 2009. Characterisation of *Bombyx mori* odorant binding proteins reveals that a general odorant binding protein discriminates between sex pheromone components. *J. Mol. Biol.* 389, 529–545.
- Zhou, J.J., 2010. Odorant-binding proteins in insects. *Vitam. Horm.* 83, 241–272.
- Zhu, F., Xu, J.J., Palli, R., Ferguson, J., Palli, S.R., 2011. Ingested RNA interference for managing the populations of the Colorado potato beetle, *Leptinotarsa decemlineata*. *Pest. Manag. Sci.* 67, 175–182.