

RESEARCH ARTICLE

Open Access

Identification and characterization of three chemosensory receptor families in the cotton bollworm *Helicoverpa armigera*

Nai-Yong Liu^{1,2†}, Wei Xu^{1†}, Alexie Papanicolaou^{1†}, Shuang-Lin Dong^{2*} and Alisha Anderson^{1*}

Abstract

Background: Chemosensory receptors including olfactory receptors (ORs), gustatory receptors (GRs) and ionotropic receptors (IRs) play a central role in sensing chemical signals and guiding insect behaviours, and are potential target genes in insect pest control. The cotton bollworm *Helicoverpa armigera* is one of the most destructive pest species that can feed on over 200 different plant species. This diversity of host plants is likely linked to a complex chemosensory system. Here we built on previous work to characterize crucial chemosensory tissues linked to environmental interactions including larval antennae, larval mouthparts and larval fat bodies, as well as male and female adult heads, male and female adult tarsi, and female abdomens.

Results: Using transcriptome sequencing, Trinity RNA-seq assemblies and extensive manual curation, we identified a total of 91 candidate chemosensory receptors (60 candidate ORs, 10 GRs and 21 IRs). Thirty-five of these candidates present full-length transcripts. First, we performed *in silico* differential expression analysis on different sequenced tissues. Further, we created extensive expression profiles using reverse transcription (RT)-PCR on a variety of adult and larval stages. We found that the expression profile of *HarmOR51* was limited to adult male antenna suggesting a role in mating that was further supported by a phylogenetic analysis clustering it into the pheromone receptor clade. *HarmOR51* in calcium imaging analysis did not show responses to either of the two *H. armigera* sex pheromone components (*Z9-16:Ald* or *Z11-16:Ald*) inviting a future detailed study. In addition, we found four novel *HarmORs* (OR1, 53, 54 and 58) that appeared to be larvae-antennal specific. Finally, our expression profiling showed that four “divergent” *HarmIRs* (IR2, 7d.1, 7d.2 and 7d.3) were expressed in both adult and larval antennae, suggesting a functional divergence from their *Drosophila* homologues.

Conclusions: This study explored three chemoreceptor superfamily genes using a curated transcriptomic approach coupled with extensive expression profiling and a more limited functional characterization. Our results have now provided an extensive resource for investigating the chemoreceptor complement of this insect pest, and meanwhile allow for targeted experiments to identify potential molecular targets for pest control and to investigate insect-plant interactions.

Keywords: *Helicoverpa armigera*, Olfactory receptor, Gustatory receptor, Ionotropic receptor, Expression profile

* Correspondence: sldong@njau.edu.cn; Alisha.Anderson@csiro.au

†Equal contributors

²Education Ministry Key Laboratory of Integrated Management of Crop Diseases and Pests, College of Plant Protection, Nanjing Agricultural University, Nanjing 210095, China

¹CSIRO Ecosystem Sciences, Black Mountain, Canberra ACT 2601, Australia

Background

The chemosensory system is critical in guiding insect feeding, mating and oviposition behaviours [1]. Hair-like chemosensory sensilla distributed over the surface of chemosensory tissues including antennae, palps, mouthparts and tarsi, are used to detect chemical signals from the external environment [2,3]. Chemosensory receptors of olfactory receptors (ORs) [4], gustatory receptors (GRs) [5] and ionotropic receptors (IRs) [6] are localized on the dendrite of chemosensory neurons housed in these sensilla. These receptors play a central role in helping insects detect chemical signals and regulate their behaviours [7], and are also important molecular targets for designing and developing new pest control strategies.

Even though *Drosophila melanogaster* is the main model insect for these genes [4-6], the availability of new genomic and transcriptomic sequences from other species is consistently extending both the phylogenetic coverage and the number of orthologs identified. The *D. melanogaster* genome hosts 60 OR genes that encode 62 ORs by alternative splicing [4]. Interestingly, insect ORs are highly diverse, share no sequence similarity and have an inverted membrane topology compared with mammalian OR genes, suggesting that insect ORs are not G-Protein Coupled Receptors (GPCRs) [8]. *Drosophila* GR genes have been classified into “CO₂” [9], “sugar” [10], “GR43a-like” [11] and “bitter” clades [12]. GRs promote insect survival by detecting nutritious compounds and help avoid toxic ones [13]. Like ORs, these receptors share no sequence similarity with vertebrate GRs and their topology is inverted compared to the classic GPCRs [14,15]. More recently, a variant sub-family of ionotropic glutamate receptors (iGluRs), the ionotropic receptor family, was identified as a new class of chemosensory receptors in *Drosophila* [6]. On the basis of their expression and sequence characteristics, *Drosophila* IRs were further distinguished into two sub-families: conserved “antennal IRs” involved in olfaction and species-specific “divergent IRs” that are expressed in other tissues including gustatory organs and may possibly be associated with gustation [16].

For this work we focused on the cotton bollworm (*Helicoverpa armigera* Hübner 1809; Lepidoptera: Noctuidae), an economically important species feeding on a wide range of host plants. *H. armigera* is one of the most polyphagous and cosmopolitan pest species with larvae that feed on numerous important cultivated crops such as cotton, peanuts, soybeans and maize. In a previous study, forty-seven OR genes were identified from *H. armigera* adult antennal transcriptomes but only one GR and 12 IR genes [17]. Only 22 full-length open reading frames were identified in this study, limiting any functional characterizations of these receptors. This previous transcriptome sequencing was restricted to adult antennae and did not consider other chemosensory

tissues. Here, we conducted additional sequencing on other potential chemosensory tissues, including taste organs, not only in adults but also in larvae. Such transcriptomic data greatly improved the description of chemosensory receptors in *H. armigera* and will also assist in defining gene models for future genome studies.

Indeed, some chemosensory gene families have a low amino acid identity with their homologues and in our experience *de-novo* gene predictors have a limited capacity in annotating them. Even though genome sequencing will eventually provide a framework for full ascertainment, an extensive manual effort is required to curate complex genes such as GRs and IRs. Moreover, in this work we first used the expression profile and phylogenetic analysis to associate each gene with putative functions, and then characterized the functions of candidate pheromone receptors by the Sf9 calcium imaging technique. In addition, we outlined a path forward for an integrated study of the insect chemosensory system that can proceed solely based on transcriptomic information.

Results

Analysis of *H. armigera* transcriptome

We prepared 10 RNA-Seq libraries which after sequencing and quality control provided us with 239,276,681 read pairs of up to 100 bp in length (Additional file 1: Table S1). Our Trinity RNA-Seq assembly produced 68,100 contigs which included redundancy due to alternative transcription or high polymorphism and non-coding regions. After an initial protein prediction using a combination of a 5th order Markov chain and the Pfam database, we created a subset of 30,111 contigs with coding potential and annotated them with controlled vocabulary terms using the Uniprot database. These open reading frames are available via the NCBI BioProject PRJNA244590 (<http://www.ncbi.nlm.nih.gov/bioproject/PRJNA244590>). The full dataset, annotations such as Gene Ontology, Enzyme Classification, and visualizations can be explored at <http://annotation.insecta-central.org/Liu2013>. After extensive manual curation, we verified or reconstructed 91 chemosensory receptors.

In silico differential expression profile of candidate *H. armigera* chemosensory receptors

To provide global gene expression profiles in different sequenced libraries, we surveyed the differential expression of all these open reading frames identified in this study. The complete result dataset is available at the CSIRO Data Portal (<http://dx.doi.org/10.4225/08/535DB2C141C8A>). As expected, all these genes (94 genes including *iGluRs*) were detected in at least one library. Most of these genes were expressed in male and/or female heads, where the antenna, proboscis and labial palp of crucial chemosensory organs, are located suggesting a functional role of these genes in olfaction or gustation. Specially, some genes appeared to be

expressed only in male and/or female heads. For example, in the GR family *HarmGR1*; in the IR family *HarmIR1*, 41a, 75p.2 and 87a; and in the OR family five of seven pheromone receptors including the newly identified pheromone receptor *HarmOR51* were detected only in male heads. In summary, from 10 sequenced libraries we found the transcriptomic support of nearly all predicted chemosensory receptor genes.

The *H. armigera* candidate olfactory receptors

We identified 57 candidate ORs with 26 full-length sequences from *H. armigera* transcriptomes of larval antennae, mouthparts and fat bodies, male and female adult heads, male and female adult tarsi, and female abdomens. This included 44 previously identified ORs [17]. We named the genes as per the conventions followed by Liu et al. [17]. Integrating all public data, we found that the previously identified *HarmOR5* was actually a gustatory receptor based on a sequence search versus NCBI's non-redundant database and a phylogenetic analysis, and therefore renamed it *HarmGR8* (Additional file 2: Table S2). We did not detect the previously identified *HarmOR28*, 33 and 37 genes in our transcriptome (Additional file 2: Table S2), so the previously published sequences were used for our phylogenetic analysis. These 60 *H. armigera* ORs were used for a phylogenetic clustering with 68 *Bombyx mori* ORs, 64 *Danaus plexippus* ORs and 69 *Heliconius melpomene* ORs (Figure 1A). In this analysis, sixteen *D. plexippus* ORs, five *H. melpomene* ORs (OR35, 36 and 38–40), and six *B. mori* ORs (OR45–48, 57 and 58) formed two monophyletic sub-families. One group of ORs formed by *BmorOR1*, 3–7 and 9, *HarmOR6*, 11, 13–16 and 51 and *DpleOR1a*, 1b, 6a and 6b was clustered with the pheromone receptor (PR) group. The novel candidate pheromone receptor *HarmOR51* shared 68.4% and 62.9% identities to *Heliothis virescens* OR14 and *HarmOR14*, respectively. Further, a phylogenetic tree from noctuid PRs and co-receptors (Orco) allowed us to classify PRs into six main sub-families: the OR6, OR11, OR13, OR14, OR15 and OR16 clades (Figure 1B). The novel *HarmOR51* was very close to the OR14 sub-family and the other six *HarmORs* were well clustered into each clade of noctuid PRs.

We used RT-PCR on adult and larval antennae to study the expression profiles of 13 newly identified *H. armigera* OR genes (*HarmOR1*, 49–60). The results showed that OR1, 53, 54 and 58 were detected only in larval antenna while the others were detected only in adult antenna, with the novel pheromone receptor *HarmOR51*, specifically expressed in the male antenna (Figure 2).

The *H. armigera* candidate gustatory receptors

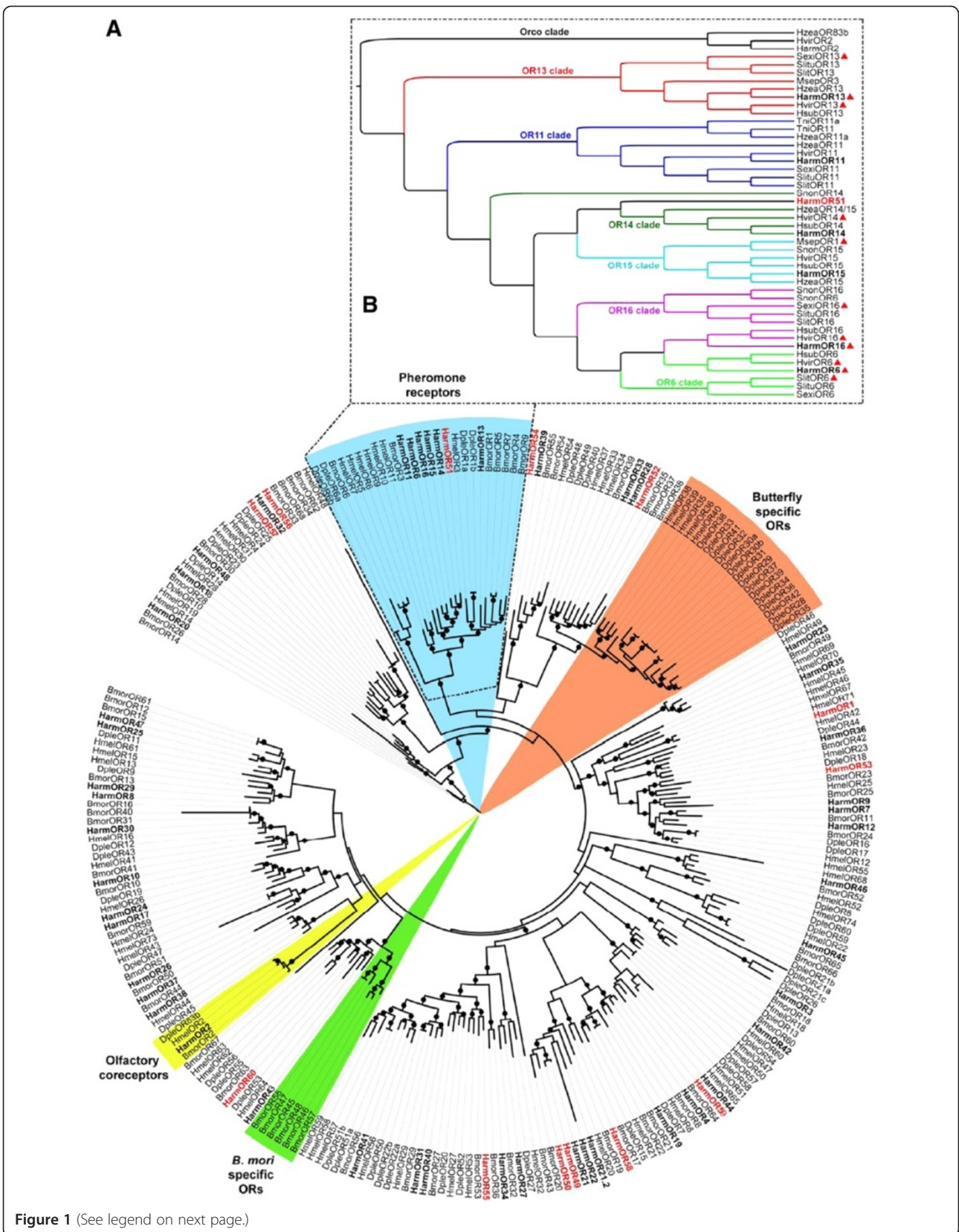
We identified nine candidate GRs from the *H. armigera* transcriptomes with *HarmGR1*–3 and *HarmGR9* as full-

length ORFs but the five others as partial sequences (Additional file 2: Table S2). We used a phylogenetic approach to name these genes. Previously reported *HarmGR1* that was not found in our transcriptome was renamed as *HarmGR10* following our identified *HarmGRs* in order to avoid a replicated name. A phylogenetic tree was built with 10 *H. armigera* GRs, 69 *B. mori* GRs and 47 *D. plexippus* GRs (Figure 3). *HarmGR1*–3 were orthologous to the silkworm “CO₂” receptors, which shared ~76 to 90% identity to *B. mori* GR1–3, respectively. *HarmGR4*–8 and GR10 were members of the insect “sugar” receptor sub-family and shared ~10 to 65% identity to *B. mori* “sugar” GRs. *HarmGR9*, one member of the “GR43a-like” receptor sub-family, was orthologous to *BmorGR9* (69% identity) and *DmelGR43a* (26% identity) receptors. None of the “bitter” receptors were found in our transcriptome sequences (Figure 3 and Additional file 2: Table S2).

Expression profiles were investigated with RT-PCR on larval antenna, larval mouthpart, male and female adult head, as well as male and female adult tarsi tissues. All of the sugar, CO₂ and GR43a-like receptor genes were detected in the male and female adult heads, *HarmGR5* and 7 were expressed in male and female adult tarsi, and *HarmGR2*, 4 and 8 were expressed only in female adult tarsi. Only *HarmGR3* was detected in larval antennae while none of GRs were detected in larval mouthparts (Figure 2).

The *H. armigera* candidate ionotropic receptors

We identified eight candidate iGluR and 20 IR genes from *H. armigera* transcriptomes including 11 of the 12 previously reported IRs [17]. *H. armigera* iGluRs were *de-novo* named using the Arabic numerals 1–8 and homologous *HarmIRs* were named based on the sequences of *D. melanogaster* and *B. mori*. One sequence did not present any similarities with reported IRs but retained their characteristic features, and thus was named *HarmIR2*. The amino acid sequences of eight candidate *HarmiGluRs* and 21 *HarmIRs* were aligned with *D. melanogaster* iGluRs. In *HarmiGluR* family with the exception of *HarmiGluR1* and 5, a conserved amino acid profile in three key residues of ligand-binding domains was observed (arginine, threonine and aspartate/glutamate). The profile was not, however, conserved for the candidate *HarmIRs*, confirming their membership to the IR sub-family rather than the iGluR one (Figure 4). Further, we identified the full-length ORFs of two candidate co-receptors: *HarmIR8a* and 25a. They shared 49% and 63% amino acid identities with *D. melanogaster* IR8a and 25a, respectively, and showed a higher amino acid identity of over 70% with the IR8a and 25a of other lepidopteran species (*B. mori*, *Cydia pomonella*, *D. plexippus*, *Manduca sexta*, *Spodoptera littoralis* and *Sesamia nonagrioides*). In addition, the full-length ORFs of a



(See figure on previous page.)

Figure 1 Phylogenetic tree of putative *H. armigera* ORs with other insect ORs. (A) The phylogenetic tree based on OR sequences from *H. armigera*, *B. mori*, *D. plexippus* and *H. melpomene*. **(B)** The phylogenetic tree based on PR sequences from 10 noctuid moths. The tree was constructed using PhyML under the JTT model of substitution with NNI topology search, based on an amino acid alignment by ClustalW2. Branch support was estimated using an approximate likelihood ratio test (Chi2) (circles: > 0.98). HarmORs are in bold and newly identified candidate HarmORs are highlighted with red letters. In Figure 1B, these PR genes with known ligands are labeled with red triangles, including *H. armigera* (this study, [18]), *H. virescens* [19,20], *Mythimna separate* [21], *S. exigua* [22], and *S. littoralis* [23]. Bmor: *B. mori*; Dple: *D. plexippus*; Harm: *H. armigera*; Hmel: *H. melpomene*; Hvir: *H. virescens*; Hsub: *H. subflexa*; Hzea: *H. zea*; Msep: *Mythimna separate*; Sexi: *S. exigua*; Slit: *S. littoralis*; Slitu: *S. litura*; Snon: *S. nonagrioides*; Tni: *Trichoplusia ni*.

Lepidoptera-specific IR87a solely expressed in adult antennae [24] and a candidate receptor IR76b involved in salt-taste coding in *D. melanogaster* [25] were also found in our data (Additional file 2: Table S2).

The phylogenetic tree of HarmiGluRs and IRs was constructed with other lepidopteran IRs and *D. melanogaster* iGluRs and IRs (Figure 5). We annotated nine new candidate HarmIR genes, five of which (IR60a, 64a, 68a, 75p.1 and 93a) were well clustered into conserved

antennal IRs and the remaining ones (IR2, 7d.1, 7d.2 and 7d.3) were classified into species-specific divergent IRs. HarmIR7d.1, 7d.2 and 7d.3 formed a sister clade to the *B. mori* IR7d clade. A new divergent HarmIR2 was not classified into any lepidopteran IR clades, but was clustered into *Drosophila* divergent IRs. In addition, two large sub-families (IR7d and IR75 clades) and a potentially specific IR group [24,26] in lepidopteran species were observed (Figure 5).

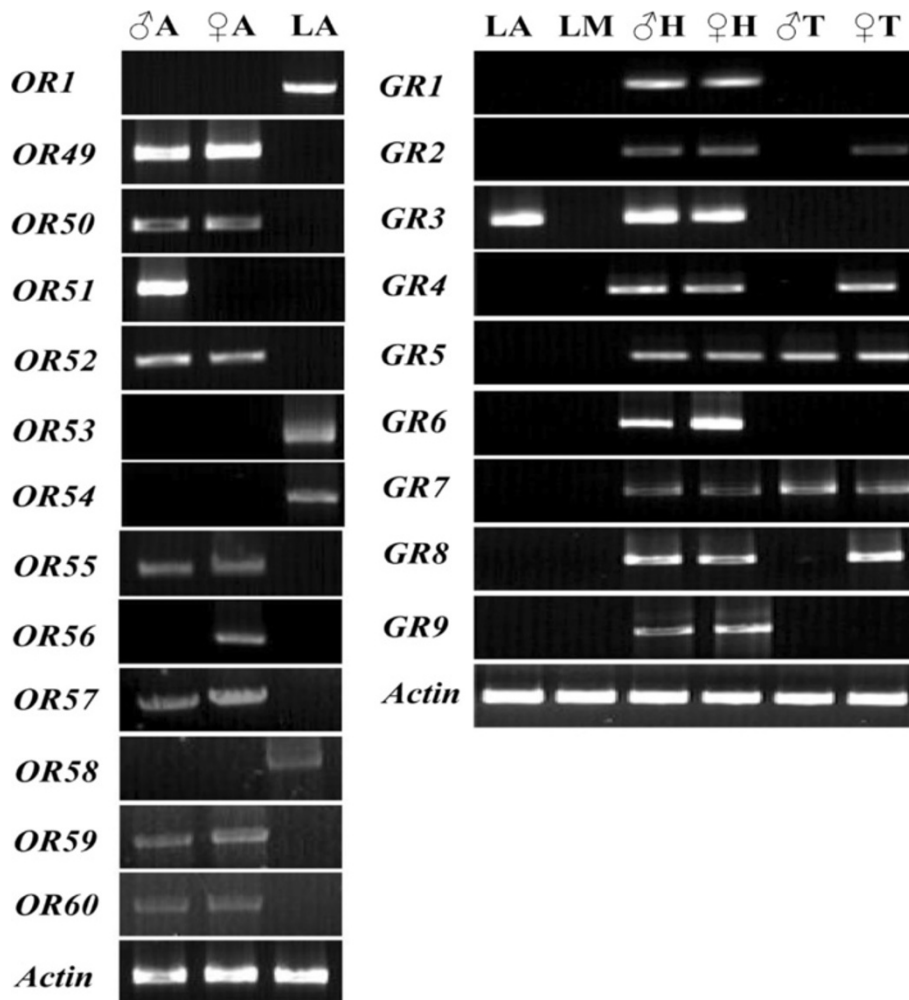
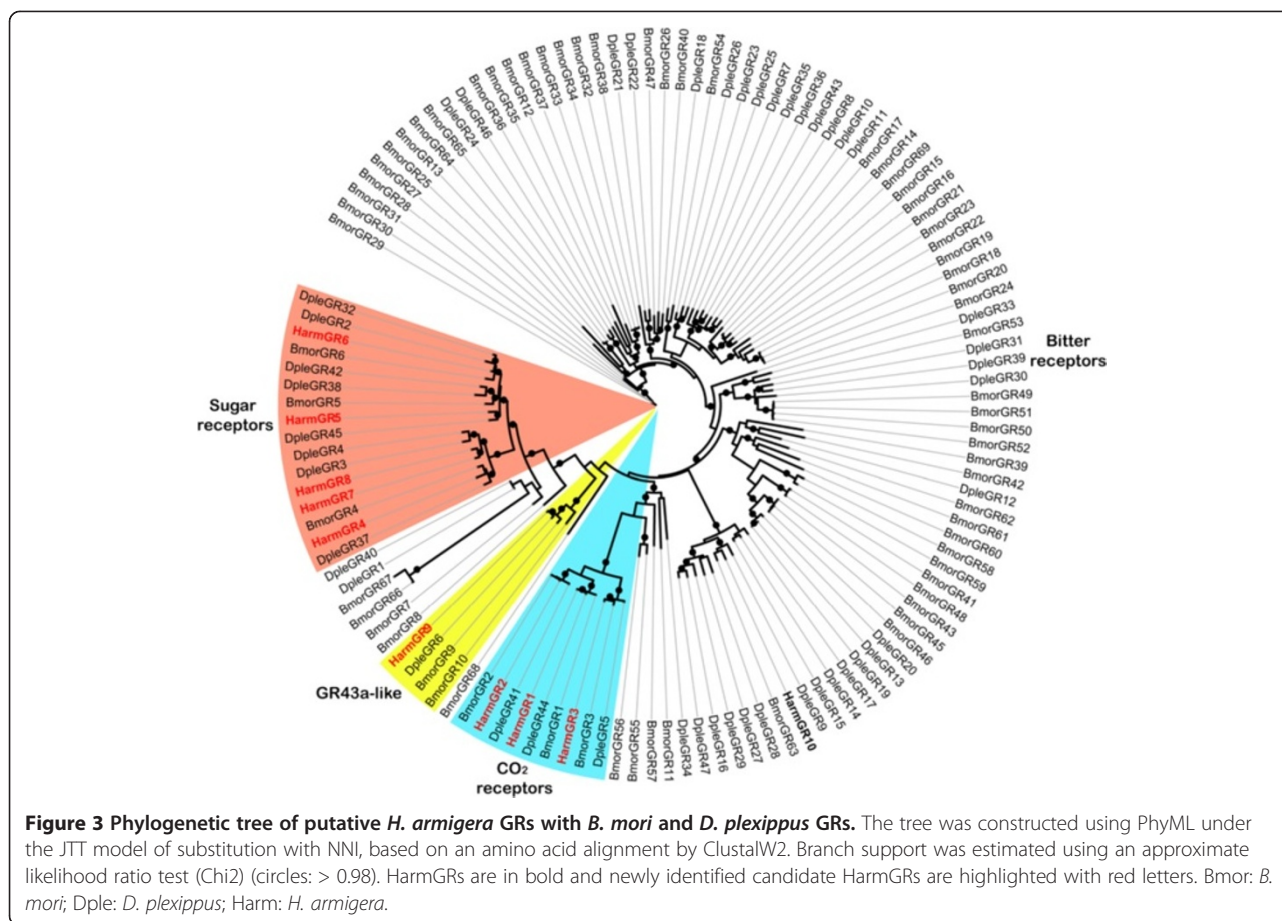


Figure 2 Adult sex- and larval tissue- specific expression of putative *H. armigera* ORs and GRs. Expression profiles of newly identified candidate HarmOR and GR genes were performed with gene-specific primers by RT-PCR. ♂A, male antenna; ♀A, female antenna; ♂H, male head; ♀H, female head; ♂T, male tarsi; ♀T, female tarsi; LA, larval antenna and LM, larval mouthpart. *Harmactin* gene was used as a control for all cDNA templates.



RT-PCR expression profiles were built for a number of adult tissues as well as larval antennae and mouthparts. In the adult tissues, the results showed that all of these *HarmIR* genes were strongly expressed in antennae of both sexes except for *HarmIR75p.1* and *75p.2* exhibiting only a low level of expression. The *HarmIR87a* gene appeared to be specifically expressed in adult antennae. Additionally, there were as many as 12 *HarmIRs* expressed in proboscises or wings. Most *HarmIR* genes were detected in larval antennae and mouthparts, with only some exhibiting no or low expression, like *HarmIR1.2*, *7d.2*, *41a*, *60a*, *75p.2* and *87a* (Figure 6).

Functional studies of candidate *H. armigera* pheromone receptors in Sf9 cells

We used Sf9 cells coupled with the calcium imaging technique [27] to assess the activity of HarmOR13 and 51 to the major sex pheromone component (Z)-11-hexadecenal (Z11-16:Ald) as well as one minor component (Z)-9-hexadecenal (Z9-16:Ald) (Figure 7). HarmOR13 showed a significant response to Z11-16:Ald but not Z9-16:Ald, as seen in previous studies using the *Xenopus* system [18]. Moreover, HarmOR13 showed an apparent dose-response to Z11-16:Ald with EC₅₀

value of 3.71×10^{-9} M. HarmOR51 did not show any responses to either Z9-16:Ald or Z11-16:Ald.

Discussion

In this work we built upon previous studies to investigate an array of tissues that play an important role in insect-plant interactions [28,29]: larval antennae, larval mouthparts, larval fat bodies, as well as female abdomens, adult tarsi and adult heads with antenna, proboscis and labial palp. Our research interest primarily focused on three crucial repertoires of chemosensory receptor super-families (ORs, GRs and IRs) due to their significance as potential target genes for designing and developing new pest control strategies and addressing the insect host-plant interactions. After the extensive sequencing and assembly using Trinity RNA-Seq, we manually curated every transcriptome so that the community now has accessed to a repertoire of 60 ORs, 10 GRs and 21 IRs. We found that *H. armigera* has more IRs than in any other lepidopteran species reported to date [30]. Further, nine *H. armigera* GRs were found in the transcriptomic data, although gustatory receptors have been reported to exhibit extremely low expression in gustatory organs [5,31]. We were also able to fully reconstruct the

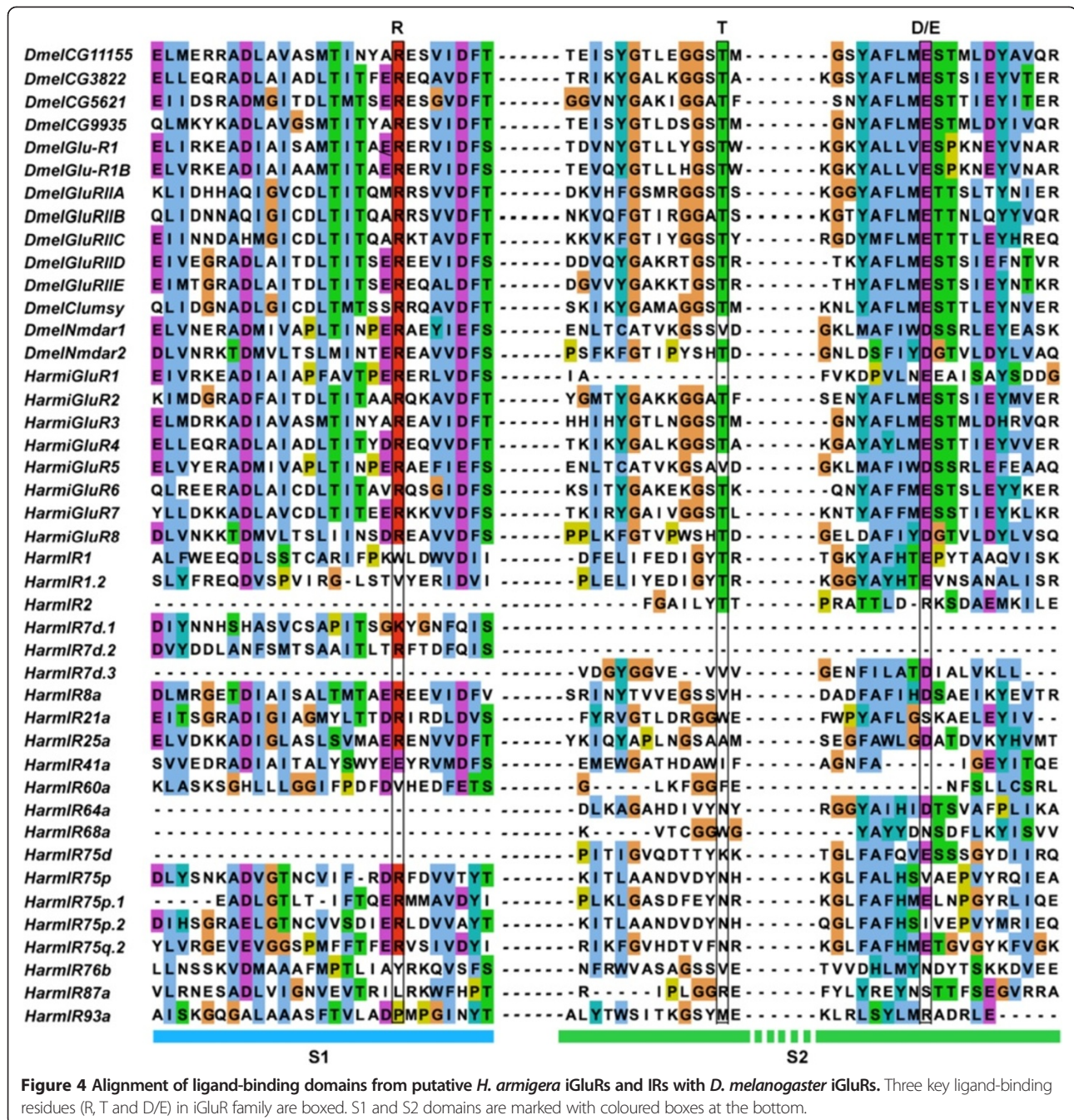
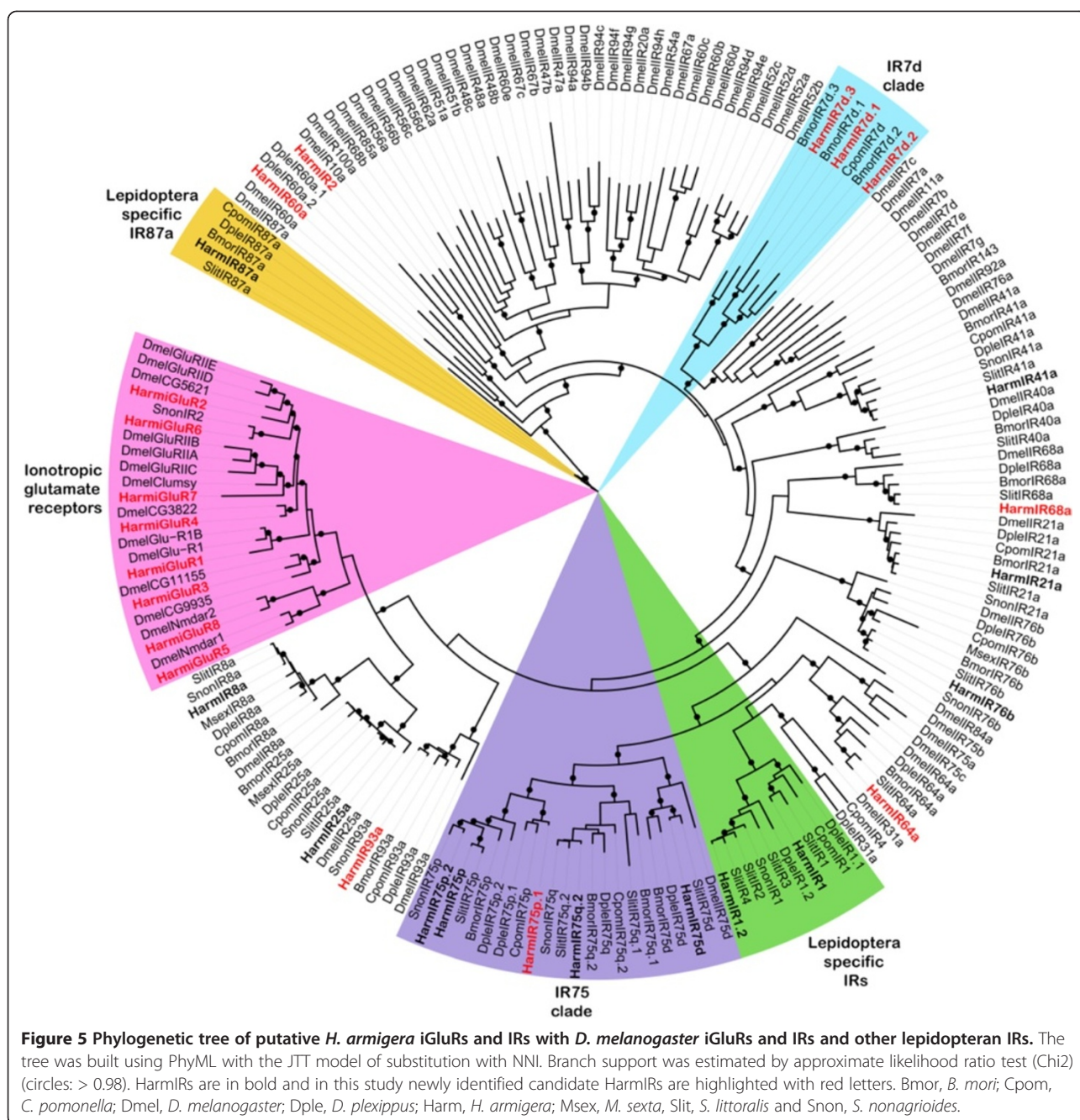


Figure 4 Alignment of ligand-binding domains from putative *H. armigera* iGluRs and IRs with *D. melanogaster* iGluRs. Three key ligand-binding residues (R, T and D/E) in iGluR family are boxed. S1 and S2 domains are marked with coloured boxes at the bottom.

transcripts of the newly predicted *H. armigera* chemosensory receptors from sequenced libraries.

In addition to verifying our transcript reconstructions, an extensive sex-, tissue- and stage-specific expression profile using RT-PCR for the newly identified ORs and GRs, as well as all the IRs allowed us to further characterize these genes. We found four larvae-specific *HarmORs* which may be involved in larvae-specific behaviours. It has been reported that *BmorOR56*, expressed in silkworm larvae, mediated larval responses

to *cis*-jasmonate that may be emitted from mulberry leaves as an attractant [28]. These larvae-specific *HarmORs* may be candidate targets for future larval control. Phylogenetic analysis showed that our newly identified *HarmOR51* was clustered into the pheromone receptor sub-family, and shared a high identity to other noctuid PRs. Thus we believe that it is possibly a new candidate of *H. armigera* pheromone receptors, leading to a total of seven pheromone receptors reported in a lepidopteran species for the first time.



Further, we used the Sf9 calcium technique to measure the binding of HarmOR13 and 51 to the *H. armigera* sex pheromone components (Z11-16:Ald or Z9-16:Ald). This system is likely to more closely resemble *in vivo* *H. armigera* olfactory receptor functions, since insect Sf9 cells are derived from a lepidopteran species *S. frugiperda*, and express a native co-receptor (Orco, a highly-conserved OR across different insect taxa) to assist HarmORs in ligand detection. We identified a strong response of HarmOR13 to Z9-16:Ald, as seen in another functional study [18]. The HarmOR51 receptor, however, did not elicit a response to

either component tested, and needs to be investigated in a comprehensive future analysis. Although our calcium imaging did not help identify the ligands for HarmOR51, it did work very well in the analysis of HarmOR13 to Z11-16:Ald, suggesting that the assay could be applied in future analyses to seek potential ligands for other ORs.

Four novel divergent *HarmIRs* were detected in both larval and adult antennae, suggesting a different function to the *Drosophila* divergent IRs that are not expressed in the antennae of any stages [16]. Our extensive expression profile showed that most *HarmIRs* were expressed in

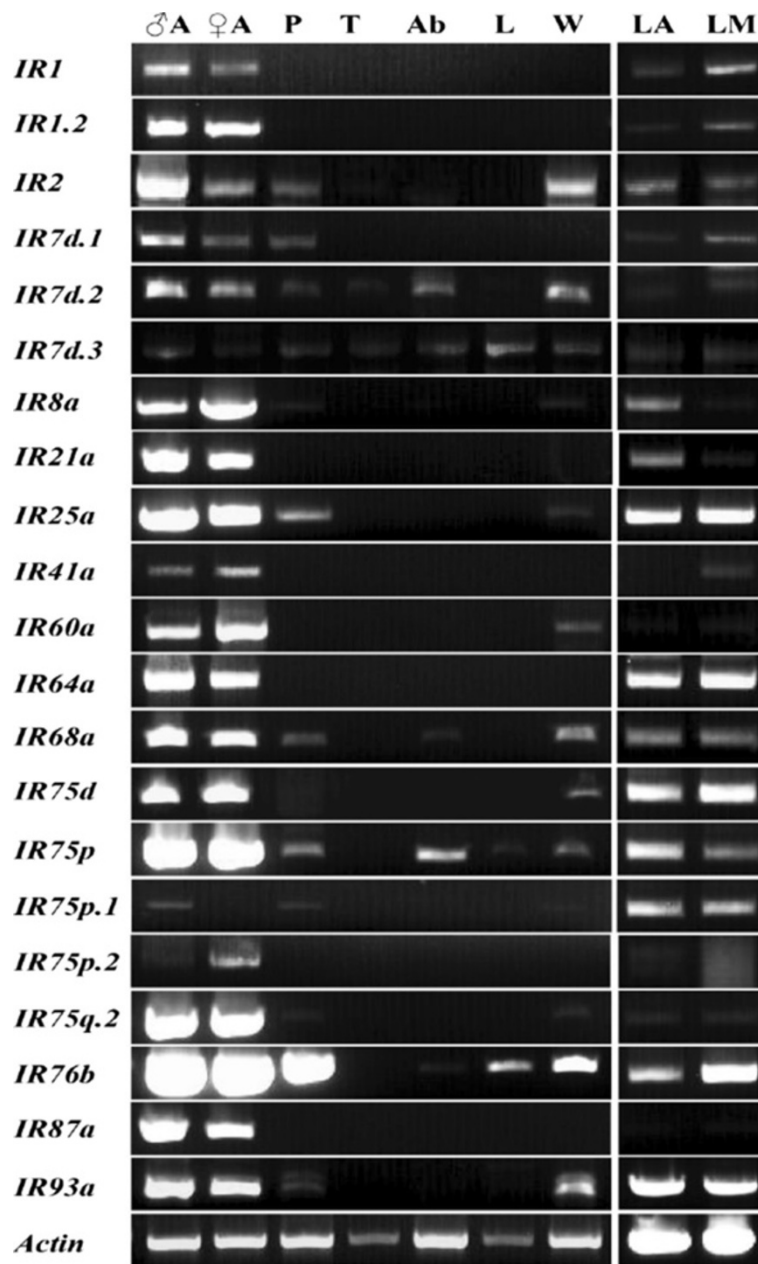


Figure 6 Adult and larval tissue-specific expression of putative *H. armigera* IRs. Expression profiles of all putative HarmIR genes were performed with gene-specific primers by RT-PCR. ♂A, male antenna; ♀A, female antenna; P, mixed proboscises; T, mixed thoraxes; Ab, mixed abdomens; L, mixed legs; W, mixed wings; LA, larval antenna and LM, larval mouthpart. *Harmactin* gene was used as a control for all cDNA templates.

antennae (at least 19 in adult and 18 in larva), providing a more extensive repertoire of antennal IRs compared to 16 of *D. melanogaster* [6,30]. Some *HarmIRs* were detected in adult proboscis and larval mouthpart, indicating that these IRs may also have a gustatory function as previously indicated in *D. melanogaster* [16] and *S. littoralis* [24].

Conclusions

In this study, we sequenced crucial chemosensory tissues, and utilized an integrated approach to probe three

chemosensory receptor repertoires in our transcriptome. We identified a total of 91 chemosensory receptor genes comprised of 60 ORs, 10 GRs and 21 IRs (an increase of 13 ORs, nine GRs and nine IRs compared to the previous work). In particular, the novel *HarmOR51*, the orthologous genes of which have been never found in any other lepidopteran species, was identified as a new potential pheromone receptor. We also carried out, for the first time, an extensive sex-, tissue- and stage- expression profile for all *H. armigera* IRs that showed a widespread

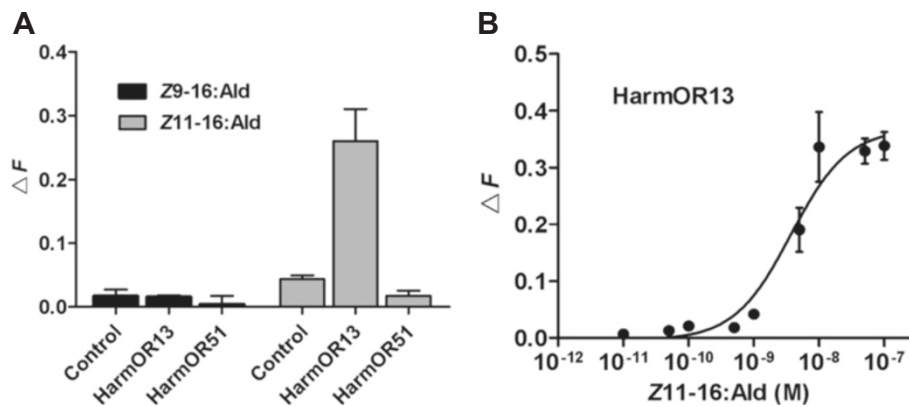


Figure 7 Calcium imaging analysis of *H. armigera* OR13 and OR51 to sex pheromone components, Z11-16:Ald and Z9-16:Ald. (A) Responses of HarmOR13 and OR51 to two sex pheromone components (10^{-7} M). HarmOR13 showed significant responses to Z11-16:Ald but not Z9-16:Ald, while HarmOR51 did not show any responses to either Z11-16:Ald or Z9-16:Ald. **(B)** Dose-dependent response curve of HarmOR13 to Z11-16:Ald. Error bars represent standard errors ($N = 3$).

expression of *HarmLRs* in non-antennal tissues, implying that in this species IRs have a more complex function. Together, this study greatly complements the information of chemosensory receptors in *H. armigera*, and shows the utility of combining transcriptomic, phylogenetic and functional assays to elucidating the function and evolution of the insect chemosensory system.

Methods

Insect rearing and tissue collection

H. armigera were fed an artificial diet in the laboratory of CSIRO under conditions described previously [32]. Approximately 200 larval antennae, 200 larval mouthparts and two larval fat bodies were dissected from fifth instar larvae. Five heads with antenna, proboscis and labial palp, 50 tarsi and two female abdomens were collected from 1 to 5-day-old male and female moths. In expression profile studies, all adult tissues were collected from 3-day-old male and female moths. All collected tissues were immediately stored in *RNAlater* (Invitrogen, USA).

Extraction of total RNA and first-strand cDNA synthesis

Total RNA was purified using RNeasy (Qiagen, USA) or RNAqueous (Ambion, USA) kits according to the manufacturer's protocol. The purified RNA was then treated with *DNaseI* (Ambion, USA) at 37°C for 30 min along with the manufacturer's protocol, quantified and qualified by NanoDrop ND-2000 (Thermo Scientific, USA) as well as 2100 Bioanalyzer (Agilent, USA).

cDNA templates were prepared from purified RNA samples of larval and adult tissues using SuperScript™ III Reverse Transcriptase (Invitrogen, USA), according to the manufacturer's manual. Briefly, the first-strand cDNA was synthesized with Oligo(dT)₂₀ primer at 65°C for 5 min, and then at 50°C for 60 min. The reaction was stopped

at 70°C for 15 min. The templates were stored at -20°C until use.

Sequencing and transcriptome analysis

The prepared RNA samples from larval antennae, larval mouthparts, larval fat bodies, male adult tarsi, female adult tarsi and female abdomens were sent to BGI-Tech for generating Paired-End Illumina TruSeq libraries and sequencing with an Illumina HiSeq 2000 platform. To help identify the chemosensory genes which are lowly expressed, we further sequenced the sample libraries of male and female adult heads that were treated by using the Duplex-Specific thermostable nuclease (DSN) enzyme (Evrogen) according to the DSN normalization protocol (Illumina). The RNA samples from male and female adult heads as well as male and female adult heads with DSN treatment were used to prepare Paired-End Illumina TruSeq libraries according to the manufacturer's protocol and sent to ACRF Biomolecular Resource Facility for Illumina sequencing service (<http://brf.anu.edu.au/>).

All RNAseq data were then pre-processed using the default settings of Just_Preprocess_My_Reads (<http://justpreprocessmyreads.sourceforge.net>) which conducted a mild quality control and trimming, pooled and assembled using Trinity-RNASeq using the default settings as per [33]. Trinity RNA-Seq is highly capable of overcoming quality and polymorphism issues due to bubble popping algorithms in each of the three modules, Inchworm, Chrysalis and Butterfly. Open reading frames (ORFs) were predicted using the TranscriptDecoder software in Trinity-RNASeq (<http://sourceforge.net/projects/transdecoder>) with the Pfam option. The raw sequence data and coding sequences of the transcriptome assembly have been deposited in the National Center for Biotechnology Information (NCBI) (accession number: PRJNA244590; <http://www.ncbi.nlm.nih.gov/bioproject/PRJNA244590>).

In silico expression profiles were generated using DEW (<http://dew.sourceforge.net/>) which is an automated pipeline that: 1) used Bowtie2 [34] to align the reads of each library against a sequence file comprised of our assembly; 2) post-processed the alignments with eXpress [35] to account for isoforms and paralogues; 3) performed Trimmed Mean Normalization using edger and estimated Fragments Per Kilobase Per Million reads (FPKM) [36]; and 4) visualized the results using ggplot2 [37] and canvasXpress (<http://canvasxpress.org/>). From these data, certain cutoffs were used to make a true/false decision on whether a gene was expressed: a gene had to have at least four RNA-Seq reads aligned against it and covering at least 0.30 of its length. Functional annotations were performed and visualised using Just_Annotate_My_Proteins (JAMp; <http://sourceforge.net/projects/jamps>) which is an automated pipeline that 1) used HHblits [38] to search against Hidden Markov Models derived from the curated Uniprot archive; 2) assigned controlled vocabulary terms (e.g. GO, KEGG etc.) linked to a Uniprot accession only if an actual experiment provides evidence (i.e. we did not use any 'inference via electronic similarity' evidence); 3) visualized them using ExtJS4 and canvasXpress. Up to 10 HHblits alignments that passed the following JAMp cutoffs were used: homology probability: $\geq 80\%$; homology e-value: $\leq 1e-10$; homology p-value: $\leq 1e-12$; HHblits score: ≥ 70 . These assemblies were then subsequently mined and curated using Geneious 5 to create a non-redundant set of the *Helicoverpa* chemosensory genes.

Identification of candidate chemosensory receptors

TBLASTN searches were carried out with previously identified *H. armigera*, *B. mori* and *Drosophila melanogaster* ORs [4,28], GRs [5,39] and IRs [6,17]. In order to verify these receptors, amino acid sequences of all identified candidate receptors were searched against NCBI non-redundant protein database using BLASTX based on the identity and similarity to orthologous genes from other insect species. These gene sequences were finally compared to previously described chemosensory receptors [17] to identify novel genes. All newly identified *H. armigera* OR, GR, as well as iGluR and IR amino acid sequences are available in GenBank as ACCESSION (Additional file 2: Table S2).

Sequence and phylogenetic analysis

Alignments of amino acid sequences were performed by ClustalW2 [40] and were visualized by Jalview 2.7 [41]. Phylogenetic trees were constructed by PhyML [42] based on Jones-Taylor-Thornton (JTT) model with Nearest Neighbour Interchange (NNI). Branch support was estimated by approximate likelihood ratio test (Chi2). In the OR data set, we selected the ORs with available genomic databases from *B. mori* [28], *D. plexippus* [43]

and *H. melpomene* [44]. Due to the small number of HarmGRs, we only selected GRs from two species, a moth *B. mori* [39] and a butterfly *D. plexippus* [43]. In the IR data set, we selected IRs from lepidopteran species including *B. mori* [16], *C. pomonella* [45], *D. plexippus* [43], *M. sexta* [46], *S. littoralis* [26] and *S. nonagrioides* [47] but also IRs and iGluRs from a model insect *D. melanogaster* [16]. Trees were viewed and edited using iTol [48,49]. Networks using the protein identity were generated using custom scripts and analyzed with CytoScape [50] and clusterMarker [51].

RT-PCR

To verify expression of candidate HarmOR, GR and IR genes identified from our transcriptome and to investigate sex- and tissue- expression profiles, RT-PCR was performed using gene-specific primers (Additional file 3: Table S3) as follows: 95°C for 3 min; 35 cycles at 95°C for 30 s, 60°C for 30 s and 72°C for 40 s; and final extension at 72°C for 7 min. *H. armigera actin* gene (GenBank accession number: X97614) was used as control to check the quality of cDNA templates. For each RT-PCR amplification, negative controls using sterile water as the template were performed. Each gene PCR reaction was repeated at least twice. PCR products were analysed using 1.5% agarose gels. The expected products of the genes randomly selected were sequenced to confirm the identity with their original sequence identified.

Calcium imaging

The sex pheromone Z11-16:Ald was purchased from Sigma-Aldrich (95% purify), and another component Z9-16:Ald was synthesized from Yick-Vic Chemicals & Pharmaceuticals (Hong Kong, China) (>95% purify). In the calcium imaging experiment, sex pheromones were dissolved in dimethyl sulfoxide (DMSO) as 1 M stock solution and stored at -20°C according to previously reported protocols [52]. For dose-dependent assays, pheromone solutions were diluted from stock solution to the desired concentration with HBSS buffer [14].

HarmOR13 and HarmOR51 ORFs were cloned with 35 cycles and an annealing temperature of 60°C using gene-specific primers (Additional file 3: Table S3). The PCR products were ligated into pBluescript vector, previously digested by EcoRV. Positive clones were sequenced. The verified plasmids were digested with KpnI and SacII for HarmOR13, and HindIII and XbaI for HarmOR51. And the digested target genes were ligated into PIB/V5-His vector, previously digested by the same enzymes. The constructed plasmids were further sequenced to confirm the orientation and sequence. Next, Sf9 cells were plated into 12-well plates and left to settle for about 20 min. Finally, the cells were transfected with 500 ng of plasmid construct PIB/HarmORs or PIB/V5-His vector (negative

control) and 3 μ L of Fugene HD transfection reagent (Promega, USA) in 100 μ L of medium per well. After 48 h post-transfection, the methods of cell treatment and functional assay were the same as previously described [14,53]. All data were analyzed by GraphPad 5.0.

Availability of supporting data

All supporting data are included as additional files. The data of transcriptome assembly were deposited in the National Center for Biotechnology Information (NCBI) (accession number: PRJNA244590; <http://www.ncbi.nlm.nih.gov/bioproject/PRJNA244590>). The transcriptome dataset and annotations can be found at <http://annotation.insectacentral.org/Liu2013>. The dataset of gene expression profiles is available at the CSIRO Data Portal (<http://dx.doi.org/10.4225/08/535DB2C141C8A>). Phylogenetic data were deposited in the CSIRO Data Portal (<http://dx.doi.org/10.4225/08/538FD6C5C9897>).

Additional files

Additional file 1: Table S1. Overview of RNASeq libraries sequenced for this study.

Additional file 2: Table S2. The identified putative *H. armigera* ORs, GRs, iGluRs and IRs.

Additional file 3: Table S3. Primers for RT-PCR experiments of *H. armigera* OR, GR and IR genes and calcium imaging analysis of *H. armigera* OR genes.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

NYL, WX and AP contributed equally to this work. This experiment was conceived and coordinated by all authors in this study. Curation and functional characterization were performed by NYL and WX. Assembly and annotation were performed by AP. NYL, WX, SLD, AA and AP drafted the manuscript. All authors read and approved the final version of the manuscript.

Acknowledgments

We are grateful to Joel Armstrong for providing *H. armigera* and Peter Campbell for help with tissue dissection. We also thank all our colleagues who provided critical feedback. Nai-Yong Liu was supported by the China Scholarship Council (CSC) under a joint PhD program with CSIRO Ecosystem Sciences, Australia. This work was supported by Commonwealth Scientific and Industrial Research Organisation (CSIRO) Julius Award (R-00094-01-005) and the CSIRO Office of the Chief Executive (OCE) Postdoctoral Fellowship (R-01479-1). During the revision of our manuscript, a related work titled "Sequence similarity and functional comparisons of pheromone receptor orthologs in two closely related *Helicoverpa* species" identified *HarmOR51* gene that was described as *HarmOR14b* in this paper [54].

Received: 31 October 2013 Accepted: 9 July 2014

Published: 15 July 2014

References

1. Leal WS: Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol* 2013, **58**:373–391.
2. Syed Z, Leal WS: Maxillary palps are broad spectrum odorant detectors in *Culex quinquefasciatus*. *Chem Senses* 2007, **32**(8):727–738.
3. Keil TA: Reconstruction and morphometry of silkworm olfactory hairs: A comparative study of sensilla trichodea on the antennae of male *Antheraea polyphemus* and *Antheraea pernyi* (Insecta, Lepidoptera). *Zoomorphology* 1984, **104**(3):147–156.
4. Clyne PJ, Warr CG, Freeman MR, Lessing D, Kim J, Carlson JR: A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. *Neuron* 1999, **22**(2):327–338.
5. Clyne PJ, Warr CG, Carlson JR: Candidate taste receptors in *Drosophila*. *Science* 2000, **287**(5459):1830–1834.
6. Benton R, Vannice KS, Gomez-Diaz C, Vosshall LB: Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* 2009, **136**(1):149–162.
7. Tunstall NE, Warr CG: Chemical communication in insects: the peripheral odour coding system of *Drosophila melanogaster*. *Adv Exp Med Biol* 2012, **739**:59–77.
8. Spehr M, Munger SD: Olfactory receptors: G protein-coupled receptors and beyond. *J Neurochem* 2009, **109**(6):1570–1583.
9. Jones WD, Cayirlioglu P, Kadow IG, Vosshall LB: Two chemosensory receptors together mediate carbon dioxide detection in *Drosophila*. *Nature* 2007, **445**(7123):86–90.
10. Slone J, Daniels J, Amrein H: Sugar receptors in *Drosophila*. *Curr Biol* 2007, **17**(20):1809–1816.
11. Sato K, Tanaka K, Touhara K: Sugar-regulated cation channel formed by an insect gustatory receptor. *Proc Natl Acad Sci U S A* 2011, **108**(28):11680–11685.
12. Lee Y, Moon SJ, Montell C: Multiple gustatory receptors required for the caffeine response in *Drosophila*. *Proc Natl Acad Sci U S A* 2009, **106**(11):4495–4500.
13. Wada-Katsumata A, Silverman J, Schal C: Changes in Taste Neurons Support the Emergence of an Adaptive Behavior in Cockroaches. *Science* 2013, **340**(6135):972–975.
14. Zhang HJ, Anderson AR, Trowell SC, Luo AR, Xiang ZH, Xia QY: Topological and functional characterization of an insect gustatory receptor. *PLoS One* 2011, **6**(8):e24111.
15. Xu W, Zhang HJ, Anderson A: A sugar gustatory receptor identified from the foregut of cotton bollworm *Helicoverpa armigera*. *J Chem Ecol* 2012, **38**(12):1513–1520.
16. Croset V, Rytz R, Cummins SF, Budd A, Brawand D, Kaessmann H, Gibson TJ, Benton R: Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet* 2010, **6**(8):e1001064.
17. Liu Y, Gu S, Zhang Y, Guo Y, Wang G: Candidate olfaction genes identified within the *Helicoverpa armigera* antennal transcriptome. *PLoS One* 2012, **7**(10):e48260.
18. Liu Y, Liu C, Lin K, Wang G: Functional specificity of sex pheromone receptors in the cotton bollworm *Helicoverpa armigera*. *PLoS One* 2013, **8**(4):e62094.
19. Wang G, Vasquez GM, Schal C, Zwiebel LJ, Gould F: Functional characterization of pheromone receptors in the tobacco budworm *Heliothis virescens*. *Insect Mol Biol* 2011, **20**(1):125–133.
20. Grosse-Wilde E, Gohl T, Bouche E, Breer H, Krieger J: Candidate pheromone receptors provide the basis for the response of distinct antennal neurons to pheromonal compounds. *Eur J Neurosci* 2007, **25**(8):2364–2373.
21. Mitsuno H, Sakurai T, Murai M, Yasuda T, Kugimiya S, Ozawa R, Toyohara H, Takabayashi J, Miyoshi H, Nishioka T: Identification of receptors of main sex-pheromone components of three Lepidopteran species. *Eur J Neurosci* 2008, **28**(5):893–902.
22. Liu C, Liu Y, Walker WB, Dong S, Wang G: Identification and functional characterization of sex pheromone receptors in beet armyworm *Spodoptera exigua* (Hübner). *Insect Biochem Mol Biol* 2013, **43**(8):747–754.
23. Montagne N, Chertemps T, Brigaud I, Francois A, Francois MC, de Fouchier A, Lucas P, Larsson MC, Jacquin-Joly E: Functional characterization of a sex pheromone receptor in the pest moth *Spodoptera littoralis* by heterologous expression in *Drosophila*. *Eur J Neurosci* 2012, **36**(5):2588–2596.
24. Olivier V, Monsempe C, Francois MC, Poivet E, Jacquin-Joly E: Candidate chemosensory ionotropic receptors in a Lepidoptera. *Insect Mol Biol* 2011, **20**(2):189–199.
25. Zhang YV, Ni J, Montell C: The Molecular Basis for Attractive Salt-Taste Coding in *Drosophila*. *Science* 2013, **340**(6138):1334–1338.
26. Poivet E, Gallot A, Montagne N, Glaser N, Legeai F, Jacquin-Joly E: A comparison of the olfactory gene repertoires of adults and larvae in the noctuid moth *Spodoptera littoralis*. *PLoS One* 2013, **8**(4):e62063.
27. Kiely A, Authier A, Kralicek AV, Warr CG, Newcomb RD: Functional analysis of a *Drosophila melanogaster* olfactory receptor expressed in Sf9 cells. *J Neurosci Methods* 2007, **159**(2):189–194.
28. Tanaka K, Uda Y, Ono Y, Nakagawa T, Suwa M, Yamaoka R, Touhara K: Highly selective tuning of a silkworm olfactory receptor to a key mulberry leaf volatile. *Curr Biol* 2009, **19**(11):881–890.

29. Briscoe AD, Macías-Muñoz A, Kozak KM, Walters JR, Yuan F, Jamie GA, Martin SH, Dasmahapatra KK, Ferguson LC, Mallet J, Jacquín-Joly E, Jiggins CD: **Female Behaviour Drives Expression and Evolution of Gustatory Receptors in Butterflies.** *PLoS Genet* 2013, **9**(7):e1003620.
30. Rytz R, Croset V, Benton R: **Ionotropic receptors (IRs): chemosensory ionotropic glutamate receptors in *Drosophila* and beyond.** *Insect Biochem Mol Biol* 2013, **43**(9):888–897.
31. Scott K, Brady R Jr, Cravchik A, Morozov P, Rzhetsky A, Zuker C, Axel R: **A chemosensory gene family encoding candidate gustatory and olfactory receptors in *Drosophila*.** *Cell* 2001, **104**(5):661–673.
32. Akhurst RJ, James W, Bird LJ, Beard C: **Resistance to the Cry1Ac delta-endotoxin of *Bacillus thuringiensis* in the cotton bollworm, *Helicoverpa armigera* (Lepidoptera: Noctuidae).** *J Econ Entomol* 2003, **96**(4):1290–1299.
33. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A: **De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis.** *Nat Protocols* 2013, **8**(8):1494–1512.
34. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nat Methods* 2012, **9**(4):357–359.
35. Roberts A, Pachter L: **Streaming fragment assignment for real-time analysis of sequencing experiments.** *Nat Methods* 2013, **10**(1):71–73.
36. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.** *Bioinformatics* 2010, **26**(1):139–140.
37. Wickham H: *ggplot2: elegant graphics for data analysis.* New York: Springer; 2009.
38. Soding J: **Protein homology detection by HMM-HMM comparison.** *Bioinformatics* 2005, **21**(7):951–960.
39. Wanner KW, Robertson HM: **The gustatory receptor family in the silkworm moth *Bombyx mori* is characterized by a large expansion of a single lineage of putative bitter receptors.** *Insect Mol Biol* 2008, **17**(6):621–629.
40. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG: **Clustal W and Clustal X version 2.0.** *Bioinformatics* 2007, **23**(21):2947–2948.
41. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ: **Jalview Version 2 - a multiple sequence alignment editor and analysis workbench.** *Bioinformatics* 2009, **25**(9):1189–1191.
42. Guindon S, Gascuel O: **A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.** *Syst Biol* 2003, **52**(5):696–704.
43. Zhan S, Merlin C, Boore JL, Reppert SM: **The monarch butterfly genome yields insights into long-distance migration.** *Cell* 2011, **147**(5):1171–1185.
44. Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley A, Nadeau NJ, Zimin AV, Hughes DST, Ferguson LC, Martin SH, Salazar C, Lewis JJ, Adler S, Ahn S-J, Baker DA, Baxter SW: **Butterfly genome reveals promiscuous exchange of mimicry adaptations among species.** *Nature* 2012, **487**(7405):94–98.
45. Bengtsson JM, Trona F, Montagne N, Anfora G, Ignell R, Witzgall P, Jacquín-Joly E: **Putative chemosensory receptors of the codling moth, *Cydia pomonella*, identified by antennal transcriptome analysis.** *PLoS One* 2012, **7**(2):e31620.
46. Grosse-Wilde E, Kuebler LS, Bucks S, Vogel H, Wicher D, Hansson BS: **Antennal transcriptome of *Manduca sexta*.** *Proc Natl Acad Sci U S A* 2011, **108**(18):7449–7454.
47. Glaser N, Gallot A, Legeai F, Montagne N, Poivet E, Harry M, Calatayud PA, Jacquín-Joly E: **Candidate chemosensory genes in the stemborer *Sesamia nonagrioides*.** *Int J Biol Sci* 2013, **9**(5):481–495.
48. Letunic I, Bork P: **Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy.** *Nucleic Acids Res* 2011, **39**(suppl 2):W475–W478.
49. Letunic I, Bork P: **Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation.** *Bioinformatics* 2007, **23**(1):127–128.
50. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res* 2003, **13**(11):2498–2504.
51. Morris JH, Apeltsin L, Newman AM, Baumbach J, Wittkop T, Su G, Bader GD, Ferrin TE: **clusterMaker: a multi-algorithm clustering plugin for Cytoscape.** *BMC Bioinformatics* 2011, **12**:436.
52. Grosse-Wilde E, Svatos A, Krieger J: **A pheromone-binding protein mediates the bombykol-induced activation of a pheromone receptor *in vitro*.** *Chem Senses* 2006, **31**(6):547–555.
53. Anderson AR, Wanner KW, Trowell SC, Warr CG, Jaquin-Joly E, Zagatti P, Robertson H, Newcomb RD: **Molecular basis of female-specific odorant responses in *Bombyx mori*.** *Insect Biochem Mol Biol* 2009, **39**(3):189–197.
54. Jiang X-J, Guo H, Di C, Yu S, Zhu L, Huang L-Q, Wang C-Z: **Sequence similarity and functional comparisons of pheromone receptor orthologs in two closely related *Helicoverpa* species.** *Insect Biochem Mol Biol* 2014, **48**:63–74.

doi:10.1186/1471-2164-15-597

Cite this article as: Liu *et al.*: Identification and characterization of three chemosensory receptor families in the cotton bollworm *Helicoverpa armigera*. *BMC Genomics* 2014 **15**:597.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

