Received: 1 July 2025 | Accepted: 16 December 2025

DOI: 10.1111/imb.70029

RESEARCH ARTICLE

Insect Molecular Biology

Genome and transcriptome-based identification and expression profiling of chemosensory gene families across developmental stages and tissues in *Sirex noctilio* (Hymenoptera: Siricidae)

Alisa Postma[1,2] | **Leandri Klynsmith**[2,3] | **Tuan A. Duong**[2] | **Jeremy Dean Allison**[3,4,5] | **Werner Smidt**[6] | **Robert M. Waterhouse**[7] | **Peter Lesny**[8] | **Jan Philip Oeyen**[9] | **Malte Petersen**[10,11] | **Sebastian Martin**[8] | **Shanlin Liu**[12] | **Xin Zhou**[13] | **Tanja Ziesmann**[14] | **Alexander Donath**[8] | **Christoph Mayer**[8] | **Bernhard Misof**[8] | **Oliver Niehuis**[15] | **Ralph S. Peters**[16] | **Lars Podsiadlowski**[8] | **Martin P. A. Coetzee**[2] | **Fourie Joubert**[1] | **Bernard Slippers**[2]

[1]Department of Biochemistry, Genetics and Microbiology, Centre for Bioinformatics and Computational Biology, University of Pretoria, Pretoria, South Africa

[2]Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa

[3]African Centre of Chemical Ecology, Innovation Africa@UP, University of Pretoria, Pretoria, South Africa

[4]Department of Zoology and Entomology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa

[5]Natural Resources Canada, Canadian Forest Service, Great Lakes Forestry Centre, Sault Ste. Marie, Ontario, Canada

[6]Division of Medical Virology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

[7]Department of Ecology and Evolution, University of Lausanne, and SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland

[8]Centre for Molecular Biodiversity Research, Leibniz Institute for the Analysis of Biodiversity Change (LIB), Bonn, Germany

[9]Norwegian Institute of Bioeconomy Research, Division of Biotechnology and Plant Health, Virusus, Bacteria and Nematodes in Forestry, Agriculture and Horticulture, Ås, Norway

[10]High Performance Computing & Analytics Lab, University of Bonn, Bonn, Germany

[11]Research IT Department, University Hospital Bonn, Bonn, Germany

[12]Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China

[13]Department of Entomology, China Agricultural University, Beijing, China

[14]Institute of Immunology, The University Medical Centre of the Johannes Gutenberg University, Mainz, Germany

[15]Department of Evolutionary Biology and Ecology, Institute of Biology I (Zoology), University of Freiburg, Freiburg, Germany

[16]Leibniz Institute for the Analysis of Biodiversity Change (LIB), Museum Koenig Bonn, Bonn, Germany

Correspondence
Bernard Slippers, Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Private Bag X20, Pretoria 0028, South Africa.
Email: bernard.slippers@fabi.up.ac.za

Abstract

The Sirex woodwasp (*Sirex noctilio*; Hymenoptera: Siricidae) is among the most destructive invasive pests affecting *Pinus* plantations worldwide. Chemosensory systems offer promising targets for pest control strategies. The identification and characterization of chemosensation genes in non-model, economically significant insects such as

[Correction added on 07 February 2026, after first online publication: The copyright line was changed.]

Funding information
DSTI NRF Centre of Excellence in Plant Health Biotechnology; Tree Protection Cooperative Programme in South Africa; Leibniz Graduate School on Genomic Biodiversity Research; German Research Foundation (DFG), Grant/Award Numbers: MI 649/16–1, NI-1387/3–1; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Numbers: 186397, 202669

Associate Editor: Michael Goodisman

S. noctilio, is an important first step towards the development of such control methods. Here we sequenced and assembled a draft genome of *S. noctilio* and performed RNA-sequencing of 15 olfactory and non-olfactory tissues to study the expression patterns of chemosensation-related genes. Specific genes, such as *SnocOR16* and *SnocSNMP1*, displayed tissue- and sex-specific expression patterns, making them particularly intriguing for their potential roles in chemosensation and oviposition. As woodwasps and their related lineages form a sister group to the majority of other Hymenoptera, including Apocrita, insights into their gene repertoires are crucial for tracing the evolutionary history of chemosensory multigene families of this ecologically and economically significant insect order. This study enhances our understanding of the molecular mechanisms underlying *S. noctilio* chemosensation, paving the way for further research in chemical ecology and the functional characterization of *S. noctilio* chemosensation genes.

KEYWORDS
chemosensory, differential gene expression, genome, Hymenoptera, invasive pest, pine, sawfly, *Sirex noctilio*, transcriptome

INTRODUCTION

Insect pests pose a significant threat to plantation forests worldwide. The Sirex woodwasp (*Sirex noctilio*; Hymenoptera: Siricidae), together with its symbiotic fungus *Amylostereum areolatum* (Russalales: Amylostereaceae), is among the most critical invasive pests of *Pinus* trees (Slippers et al., 2012). Native to Eurasia and North Africa, *Sirex* has spread globally, becoming one of the most economically significant pests of softwood plantations in invaded regions, particularly in the Southern Hemisphere (Hurley et al., 2007).

The *Sirex-Amylostereum* complex is primarily managed through biological control and silvicultural practices. However, these strategies are not always fully effective (Slippers et al., 2012). Chemosensation has emerged as a promising target for pest control strategies, although much of the research has been focused on model organisms (Lizana et al., 2022). Therefore, identifying and characterizing genes involved in chemosensation in non-model, economically significant insects such as *S. noctilio* is an important step towards the development of innovative and effective control strategies.

Sirex noctilio belongs to the paraphyletic assemblage of Hymenoptera known as Symphyta, which lack the wasp waist characteristic found in the monophyletic and highly diverse Apocrita (Peters et al., 2017). The Apocrita clade includes ants, (social) wasps and bees. Although less species-rich than Apocrita, Symphyta species fill a wide range of ecological roles, including defoliation (e.g., *Athalia rosae*), stem-feeding (e.g., *Cephus cinctus*) and insect parasitoidism (e.g., *Orussus abietinus*) (Oeyen et al., 2020; Robertson, Waterhouse, et al., 2018). *S. noctilio* adopted an endophytic larval lifestyle, with its larvae developing within the woody trunk of *Pinus* species. These ecological adaptations are underpinned by specialized and unique chemosensory systems.

Classical chemical ecology techniques have been utilized to investigate host and mate location in *S. noctilio*, resulting in the development of a host kairomone lure for population monitoring (Allison, 2021). Cuticular lipids, acting as short-range contact sex pheromones to induce copulation, have been identified (Böröczky et al., 2009). However, the existence of long-range sex and aggregation pheromones remains uncertain. Putative male-produced pheromones have been identified and are thought to mediate the aggregation of male woodwasps (so-called leks), potentially aiding female wasps in locating these male leks (Guignard et al., 2021). Olfactometer bioassays have shown that *S. noctilio* females are attracted to volatiles emitted by its fungal symbiont *A. areolatum*, though the specific chemical compounds responsible for this attraction remain unknown (Fernández Ajó et al., 2015). Hayes et al. (2015) demonstrated chemosensory capabilities in *S. noctilio* ovipositors using electrophysiological analyses. Despite these advancements, knowledge of the molecular mechanisms underlying chemosensation of this insect species remains limited.

To address gaps in our understanding of the chemical and behavioural ecology of *S. noctilio*, the genomic and transcriptomic components underlying these mechanisms require further exploration. A previous study examined the chemosensory gene repertoires in the antennal transcriptomes of *S. noctilio* and *Sirex nitobei* (Hymenoptera: Siricidae) (Guo et al., 2021). However, this research focused exclusively on a single tissue type, which may have resulted in incomplete and biased identification of candidate genes and their expression profiles.

The present study aims to comprehensively characterize the chemosensory gene repertoire of *S. noctilio* by integrating both genomic and transcriptomic data. Differences in chemosensory gene expression patterns were analysed across tissue types and between sexes. Based on the analysis of 15 distinct transcriptomes, we identified and annotated six major chemosensory gene families—chemosensory proteins (CSPs), gustatory receptors (GRs), ionotropic receptors (IRs), odorant-binding proteins (OBPs), odorant receptors (ORs) and sensory neuron membrane proteins (SNMPs)—from the *S. noctilio* genome and conducted phylogenetic analyses to compare these genes with orthologues from other insect species. These efforts enabled us to generate detailed expression profiles and investigate the differential expression patterns of these genes across developmental stages, tissues and sexes.

EXPERIMENTAL PROCEDURES

Whole genome sequencing

High molecular weight DNA was extracted from whole-body tissue (after removing the gut) of haploid male individuals from the Cape region in South Africa, using a modified phenol-chloroform DNA extraction protocol (File S1). Two paired-end libraries (read length 150 bp, insert size 250 bp and read length 100 bp, insert size 300 bp) and one mate-pair library (read length 50 bp, insert size 2 kbp) were constructed and sequenced on the Illumina HiSeq platform by Fasteris, Life Science Genesupport SA (Plan-Les-Quates, Geneva, Switzerland). Additionally, two mate-pair libraries with insert sizes of 3 and 8 kbp were constructed and sequenced on the Illumina HiSeq 2000 platform by BGI-Shenzhen (Shenzhen, Guangdong, China). For these libraries, sequencing was performed on a single pooled sample of genomic DNA extracted from the whole bodies of three adult males and two adult females collected in the eastern United States (Pennsylvania and New York) using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA), following the manufacturer's instructions. All DNA-sequencing details are summarized in Table S2.

Transcriptome sequencing

For genome annotation, whole-body transcriptome sequencing was performed using all tissues except the gut (File S4) from male individuals (all transcriptome sample information is summarized in Table S1). For chemosensory transcriptome sequencing, various tissues—including antennae, heads (without antennae), legs, wings, ovipositors and bodies (abdomen without gut)—were dissected from adult male and female woodwasps, respectively. Twenty-five individual adult woodwasps were collected, from which the different tissues were dissected. Four individual pupae were collected, and each pupa was dissected into head (including antennae) and body sections (remaining pupal tissue). All dissected tissue samples were immediately flash-frozen and stored at −80°C until further processing. Three replicates were prepared per tissue type for both adult and pupa samples. A negative extraction control sample as well as a reagent blank were included to check for contamination in the reagents and elution buffer.

High-quality RNA (RIN values >7) was extracted from all tissues except the gut using a modified TRIzol extraction method (File S2) (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Contamination-control measures were taken to ensure high-quality RNA extraction. Sterilizing wash procedures were followed and all dissection and extraction work was done in RNase-free environments using RNase-free tools (File S3). RNA purity and integrity were assessed with a ND-100 spectrophotometer (Nanodrop, Delaware, USA) and gel electrophoresis, respectively. RNA quantity was assessed using a Qubit 2.0 Fluorometer (Invitrogen, Waltham, Massachusetts, USA) with the Qubit RNA BR Assay Kit (Invitrogen) following the manufacturer's instructions. Additionally, RNA samples were treated with DNase I (RNase-free, 1 U/μL Thermo Fisher Scientific) according to the manufacturer's protocol, including all recommended steps. After DNAse treatment, RNA samples were purified using the RNeasy Mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

Transcriptome libraries were sequenced using the Illumina HiSeq mRNA-SEQ platform (Illumina, San Diego, California, USA). RNA sequencing of the 45 RNA samples (15 male adult samples, 18 female adult samples, 12 pupa samples) was performed by Macrogen HQ (Seoul, Republic of Korea). Library construction was carried out using the TruSeq Stranded mRNA LT Sample Prep Kit (Illumina) and sequencing was conducted on the Illumina NovaSeq 6000 platform (Illumina) to generate paired-end 101 bp stranded mRNA reads. A summary of all RNA-sequencing information is provided in Table S1. Samples were collected, prepared and sequenced to minimize batch effects. Adult wasps were collected as they emerged from the trunk. Fresh tissue samples were dissected and flash-frozen. RNA extractions were performed by one scientist using the same reagents and equipment. Library preparation and sequencing were performed by one scientist on one flow cell.

Genome assembly and annotation

The quality of all raw genome sequencing data (Table S2) was assessed using FastQC version 0.11 (Andrews, 2010) and MultiQC version 1.6 (Ewels et al., 2016). Quality trimming and filtering were performed using Trimmomatic version 0.36 (Bolger et al., 2014). Paired-end assemblies were generated using all paired-end data and the VelvetOptimiser script version 2.2.5 (https://github.com/tseemann/VelvetOptimiser) (Zerbino & Birney, 2008). The contig assembly obtained with paired-end data was then scaffolded with mate-pair data using the SSPACE BASIC script (Boetzer et al., 2011).

RNA-sequencing reads generated for improved genome annotation were mapped to the genome assembly using Tophat2 version 2.1.0 (Kim et al., 2013). These transcriptomic reads mapped back to the genome assembly at a rate greater than 90% (Table S6), which indicated that the assembly is largely complete and that most of the expressed genes were present. The mapped RNA-Seq data served as supporting evidence for genome annotation using BRAKER1 version 1.3.1 (Hoff et al., 2016). The AUGUSTUS gene prediction tool version 3.2.2 (Stanke et al., 2008) was employed for gene prediction. MAKER version 2.31_r1081 (Campbell et al., 2014) was used to integrate both intrinsic and extrinsic annotation data, generating the official gene set. The completeness of the genome assembly and annotation was assessed using the Benchmarking Universal Single-Copy Orthologs (BUSCO) tool version 3.0.2 (Simão et al., 2015).

Identification and characterization of *S. noctilio* chemosensation genes

Amino acid sequences of previously reported Hymenoptera CSPs, GRs, IRs, OBPs, ORs and SNMPs (Table S3) were used as queries to search

against the predicted proteome of *S. noctilio* using BLASTp with an E-value threshold of 1e-5 (BLAST+ software suite 2.8.0) (Camacho et al., 2009). Additionally, the same query sequences were used to search against the *S. noctilio* genome assembly via tBLASTn, with an *E*-value threshold of 1e-3 (BLAST+ software suite 2.8.0). Search results were filtered based on bit scores and alignment length. Candidate chemosensation genes were manually curated using the web-based genomic annotation platform Apollo version 2.0.6 (Lee et al., 2013). Gene models were curated based on a combination of intrinsic (start and stop sites, exon-intron structure) and extrinsic evidence (RNA-Seq data, alignment with homologues and/or orthologues and functional domains). Genes located near the ends of the scaffolds were checked for completeness by performing multiple sequence alignment (MAFFT version 7.487) of the predicted gene models with closely related sequences available in public databases. The final numbers of *S. noctilio* chemosensation genes were consistent with those previously reported for other Hymenoptera species (Table S5). Predicted *S. noctilio* proteins identified as potential chemoreceptors (GRs, IRs, ORs) were screened for target gene family-specific transmembrane and receptor domains (InterPro identifiers IPR004117, IPR009318 and IPR013604) using TMHMM version 2.0 (Krogh et al., 2001) and InterProScan version 61.0 (Jones et al., 2014), respectively. CSPs and OBPs were screened for the presence of binding protein domains (IPR006170) and signal peptides using InterProScan version 61.0 and SignalP version 4.0 (Petersen et al., 2011).

Phylogenetic analysis of insect chemosensation proteins

Phylogenetic analyses were conducted using the amino acid sequences of chemosensation proteins encoded by the *S. noctilio* genome, alongside chemosensory proteins from various other insect species (Table S4). Multiple amino acid sequence alignments (MSAs) were generated with MAFFT version 7.487 with the high-accuracy E-INS-I algorithm (Katoh & Standley, 2013). The MSAs were further refined with trimAl version 1.3 (Capella-Gutiérrez et al., 2009) (with 'gappyout' settings for GRs, OBPs and ORs and compared to the 'strict' setting for IRs). The MSA trimming parameters were chosen based on those used for these gene families in similar studies (Eyun et al., 2017; Robertson, Baits, et al., 2018).

Best-fitting models under the Bayesian information criterion (BIC) selection criterion were identified for chemoreceptors and CSPs, respectively, using ModelFinder (Kalyaanamoorthy et al., 2017) in IQ-TREE2 version 2.0.8 (Nguyen et al., 2015).

Maximum likelihood trees were inferred using IQ-TREE2 applying the best-fitting models. Branch support values were calculated from 1000 bootstrap replicates using an ultrafast bootstrap approximation. Phylogenetic trees were visualized and edited using the R package ggtree version 3.12.0 (Yu et al., 2016). Orthology relationships between chemosensation genes across different insect species were determined using Orthofinder version 2.3.8 with an inflation factor of 1.5 (Emms & Kelly, 2019).

Transcriptome and differential expression analysis of chemosensation genes

High-quality RNA reads were mapped to the draft reference genome assembly of *S. noctilio* using HISAT2 version 2.1.0 (Kim et al., 2015). RNA-sequencing read alignment to the reference genome was quantified using featureCounts (R version 3.6.0, subread version 1.5.2, http://subread.sourceforge.net) (Liao et al., 2014). Transcript abundance in TPM (transcript per million) was estimated using StringTie version 2 (Pertea et al., 2015) and was used for qualitative assessment of overall gene expression results, but not for comparative statistical analyses.

Differential expression and statistical analyses were conducted using DESeq2 version 3.2.0 (Love et al., 2014). Normalized count data were used for comparing overall gene expression patterns. These data were transformed using a variance stabilizing method, and hierarchical clustering of chemosensation genes was performed using a k-means approach in DESeq2. Principal component analysis (PCA) plots were generated using the plotPCA function in DESeq2. Heatmaps were created with the pheatmap R package (https://CRAN.Rproject.org/package=pheatmap).

To identify differentially expressed genes, pairwise comparisons were performed between conditions. Both simple (design = ~condition) and complex (design = ~condition + sex + tissue + lifestage) experimental designs as well as various contrasts were tested and applied during DESeq2 analyses. Log fold change (|log2FoldChange| ≥ 1) and adjusted *p*-values (adj. $p < 0.05$) were used as threshold criteria to identify significantly differentially expressed genes. False discovery rate (FDR) was controlled using the Benjamini-Hochberg method. To analyse effect-size distributions, a volcano plot was generated using DESeq2.

To further investigate the co-expression and co-regulation of the chemosensory genes studied, Weighted Gene Co-expression Network Analysis (WGCNA) (Langfelder & Horvath, 2008; Langfelder & Horvath, 2012) and Gene Set Variation Analysis (GSVA) (Hänzelmann et al., 2013) were performed. For co-expression analyses, the WGCNA Bioconductor R package v3.5.1 was used. Variance Stabilizing Transformation (VST) data from DESeq2 were used as input with the 'signed hybrid' network type, a maximum block size of 9500 and the soft-threshold power (β) set at 9. The per-sample enrichment of particular modules was analysed using the GSVA Bioconductor R package version 2.3.1 with default parameters. Results obtained from these analyses were plotted using the R ggplot2 and pheatmap packages.

RESULTS

Genome sequencing, assembly and annotation

In this study, a draft genome assembly of *S. noctilio* was generated. The assembled genome consisted of 6215 contigs with a total size of 185 Mbp and an N50 length of 825 kbp. The assembly has a BUSCO

completeness score of 95.4%, and 20,629 genes were predicted during the annotation process. These metrics are compared to other available 'Symphyta' genome assemblies in Table S4. A detailed summary of the genome sequencing data is provided in Table S2.

Transcriptome sequencing

For differential expression analysis, transcriptomes were generated from 15 samples, including six tissues from female wasps, five tissues from male wasps and two tissues from female and male pupae, respectively. Each sequenced library produced over 51 million paired-end, stranded reads. The percentage of bases with a minimum quality score of Q30 exceeded 92% (Table S1) for all sequences. The raw reads have been deposited at the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database under BioProject ID PRJNA982519. A detailed summary of the RNA-sequencing data is provided in Table S1.

Identification, characterization and phylogenetic analysis of *S. noctilio* chemosensation genes

A total of 129 chemosensory genes, representing six chemosensory gene families, were identified in the genome of *S. noctilio*. Phylogenetic analyses, which included the predicted *S. noctilio* chemosensation amino acid sequences alongside proteins from selected Hymenoptera species (Table S4), revealed that most *S. noctilio* chemosensation genes originated from single-gene lineages at the base of chemosensory subfamily clades.

Forty-five ORs were identified in the genome of *S. noctilio* (SnocOR1–44, SnocORco) (Figure 1). Predicted protein domain analysis confirmed the presence of insect olfactory receptor protein domains (IPR004117) in all *S. noctilio* ORs. The *S. noctilio* genome included one orthologue in the Odorant Receptor Co-receptor clade (SnocORco), one orthologue in the 9-exon superfamily clade (SnocOR16) and one orthologue in the subfamily containing AmelOR11 (SnocOR24). Expansions in OR gene numbers were noted

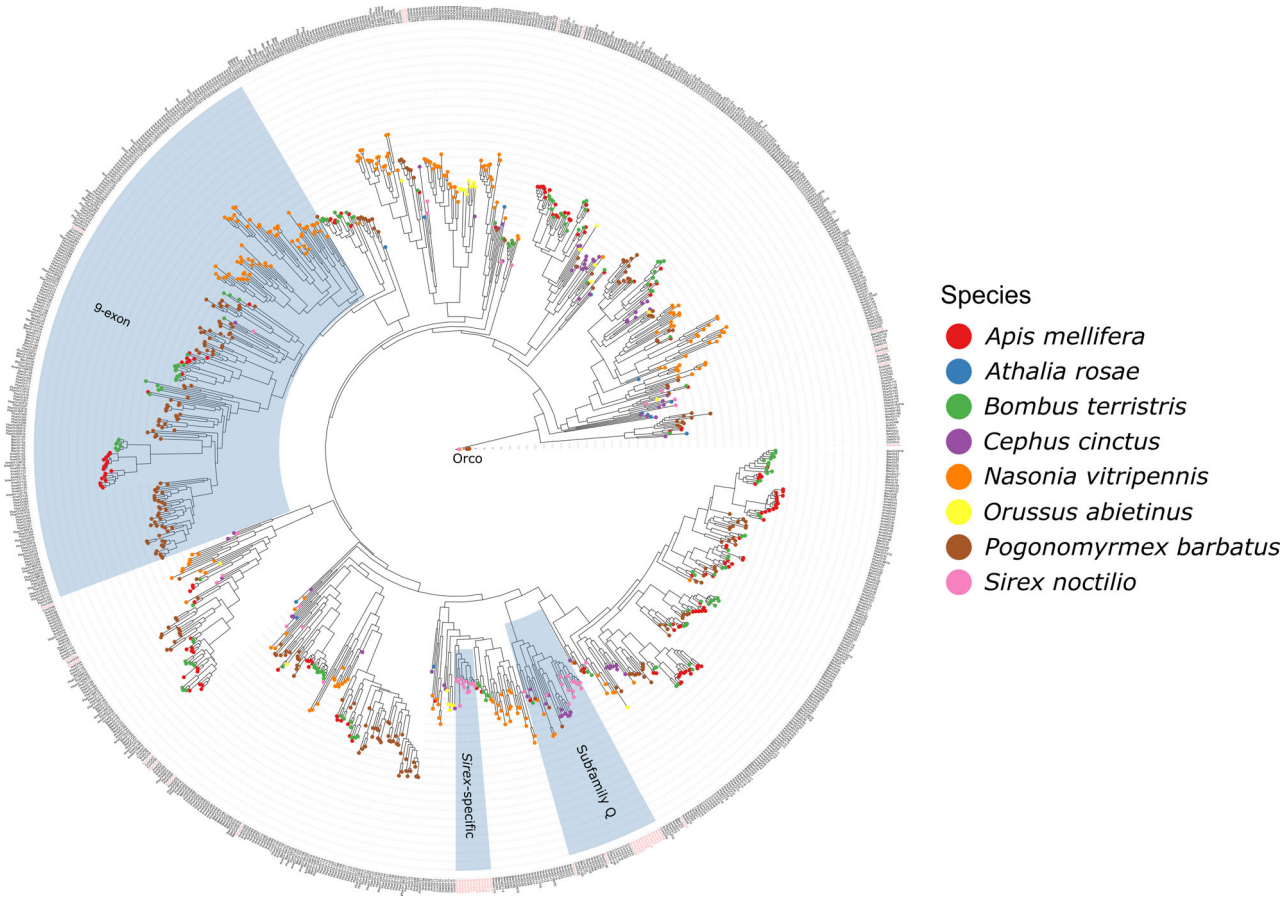

FIGURE 1 Phylogenetic tree depicting the evolutionary relationship between the 45 identified *Sirex noctilio* odorant receptor (OR) genes and OR genes from seven other Hymenoptera species. The genes from each species are depicted in colours as described in the key on the right. *S. noctilio* genes are highlighted in bold fontface and a colour matching the species key. Details of organisms and sequences used for phylogenetic analyses are listed in Table S3. *Sirex noctilio* had one orthologue in the Odorant Receptor Co-receptor clade (SnocORco) and one orthologue in the 9-exon superfamily clade (SnocOR16). These clades are highlighted with blue shading. Subfamily Q and a *Sirex*-specific clade are also highlighted. This figure is also available on the figshare repository: 10.6084/m9.figshare.30311611.

within the 9-exon superfamily and the Q subfamily. Additionally, a *Sirex*-specific expansion of OR genes was observed (Figure 1).

Twelve GRs were identified in the genome of *S. noctilio* (SnocGR1–12) (Figures S2 and S3). All identified GRs contained seven transmembrane chemoreceptor (IPR013604) or GR (IPR009318) domains. Phylogenetic analyses, incorporating the *S. noctilio* GR proteins and selected insect GR proteins, revealed that *S. noctilio* GRs typically form single-gene lineages at the base of subfamily clades. This set includes putative sugar receptors (SnocGR5 and SnocGR6), fructose receptors (SnocGR3, 4, 7, 9 and 10) and a carbon dioxide receptor (SnocGR11).

Forty-eight IRs were identified in the genome of *S. noctilio* (SnocIR1–54) (Figures S4 and S5). Protein domains associated with IRs (IPR019594, IPR001320, IPR005055 and IPR036682) were present in all identified IRs. Orthology and phylogenetic analyses revealed that *S. noctilio* possesses single orthologues of the phylogenetically conserved ionotropic co-receptors, IR8a, IR25a and IR76b (SnocIR8a, SnocIR45 and SnocIR9). Additionally, *S. noctilio* contains orthologues (SnocIR28, SnocIR54 and SnocIR31) of three *Drosophila melanogaster* genes known to be involved in thermo- and hygrosensation (DmelIR21a, DmelIR68a and DmelIR93a).

Fourteen OBPs were identified in the genome of *S. noctilio* (SnocOBP1–14) (Figures S6 and S7). Protein domains associated with pheromone or general CSPs (IPR036728, IPR006170) are present in all identified OBPs. Seven of these OBPs (SnocOBP4, 5, 8, 10–13) are classic OBPs, each containing six cysteine residues. Another six genes (SnocOBP1–3, 7, 9 and 14) were classified as non-classical OBPs, as they contain more than six cysteine residues. Signal peptides were predicted for all *S. noctilio* OBPs, with potential cleavage sites between residues 17 and 24. Phylogenetic analysis placed *S. noctilio* OBPs primarily at the base of subfamily clades. One *S. noctilio* orthologue (SnocOBP1) was found in the wasp-specific clade, but no orthologues were identified in the bee or ant-specific clades.

Additionally, nine chemosensory proteins (CSPs) were identified in the *S. noctilio* genome. All identified CSPs contain protein domains associated with insect OBPs (IPR005055, IPR036682) and the conserved four cysteine motif. Furthermore, one sensory neuron membrane protein 1 (SNMP1) was identified, which contains a CD36 family domain (IPR002159). This gene is predicted to have eight exons and two transmembrane domains.

Transcriptome and differential expression analysis of chemosensation genes

Global and chemosensation gene expression patterns were analysed to identify unique features of chemosensation gene expression (Figure 2). PCA plots indicate that samples from the same experimental groups (developmental stages, sexes and tissues) cluster together. This provided further evidence that batch effects played a negligible role in this dataset. Clustering of global gene expression data revealed clear separation based on developmental stages, whereas clustering of chemosensation gene specific expression data showed distinct separation by tissue types. Antennal samples formed unique clusters, with this distinction being more pronounced in the chemosensation gene dataset. Notably, pupal tissues clustered more closely together in the chemosensation gene dataset compared to global gene expression patterns. Overall, developmental stage and tissue type had a greater influence on gene expression patterns than sex differences. Given the clustering patterns, the use of standard DESeq2 thresholds was accepted as appropriate. To analyse effect-size distributions, a volcano plot was generated (Figure S10). This plot effectively displayed the full effect-size distribution for all tested genes and confirmed that the selection criteria applied were appropriately balanced to identify genes exhibiting high statistical significance and biologically meaningful effect sizes.

OR gene expression was generally higher in antennae compared to other tissues (Figure 3, Figure S1). Clustering analysis of normalized count data across all tissues, sexes and developmental stages revealed that all but three ORs exhibited elevated expression in antennae. Some ORs also showed sex-specific expression patterns: seven genes (SnocOR6, 9, 22–24, 28 and 39) were significantly upregulated in female antennae, while one (SnocOR16) was upregulated in male antennae ($|\log2FoldChange| \geq 1$, adj. $p < 0.05$). Additionally, SnocOR21 and SnocOR24 showed increased expression in ovipositors. The *S. noctilio* SNMP1 gene was highly expressed in adult antennae and male legs, with significantly higher expression levels in male legs compared to female legs ($|\log2FoldChange| \geq 1$, adj. $p < 0.05$).

Further analyses were performed to investigate whether the chemosensory genes studied here are co-expressed and co-regulated. Clusters (modules) of correlated genes were identified through WGCNA of the differential gene expression data. Genes that are highly connected to other genes in the same module are called hub genes. The most prominent modules found are shown in a hub-plot in Figure 4. The blue, brown, darkgrey, darkturqoise, magenta, red and turquoise modules contained the greatest number of hub genes and were selected for further investigation. The magenta module contained the greatest number of chemosensory genes (46 of the total 129 chemosensory genes reported here), as well as the largest proportion of chemosensory genes displaying high correlation with the eigengene of the module (Table S7).

GSVA was performed to study per-sample enrichment scores across all samples for the modules identified by WGCNA. The GSVA enrichment plot (Figure S8) provided additional evidence that the antennae exhibited enriched gene expression patterns compared to other tissues studied. In order to obtain a high-level perspective of the modules and potential pathways activated in certain samples, the GSVA data were further visualized using a heatmap (Figure S9). The magenta module exhibited a unique expression pattern across samples. Several chemosensory genes found to be both significantly differentially expressed and orthologous or phylogenetically related to key insect chemosensory genes were clustered into the magenta module. These genes include SnocOR9, SnocOR14, SnocOR16, SnocOR20, SnocOR21, SnocIR8a, SnocOBP3 and SnocSNMP1. Many of the non-chemosensory genes present in the magenta module were predicted to be transcription factors. Other predicted functions

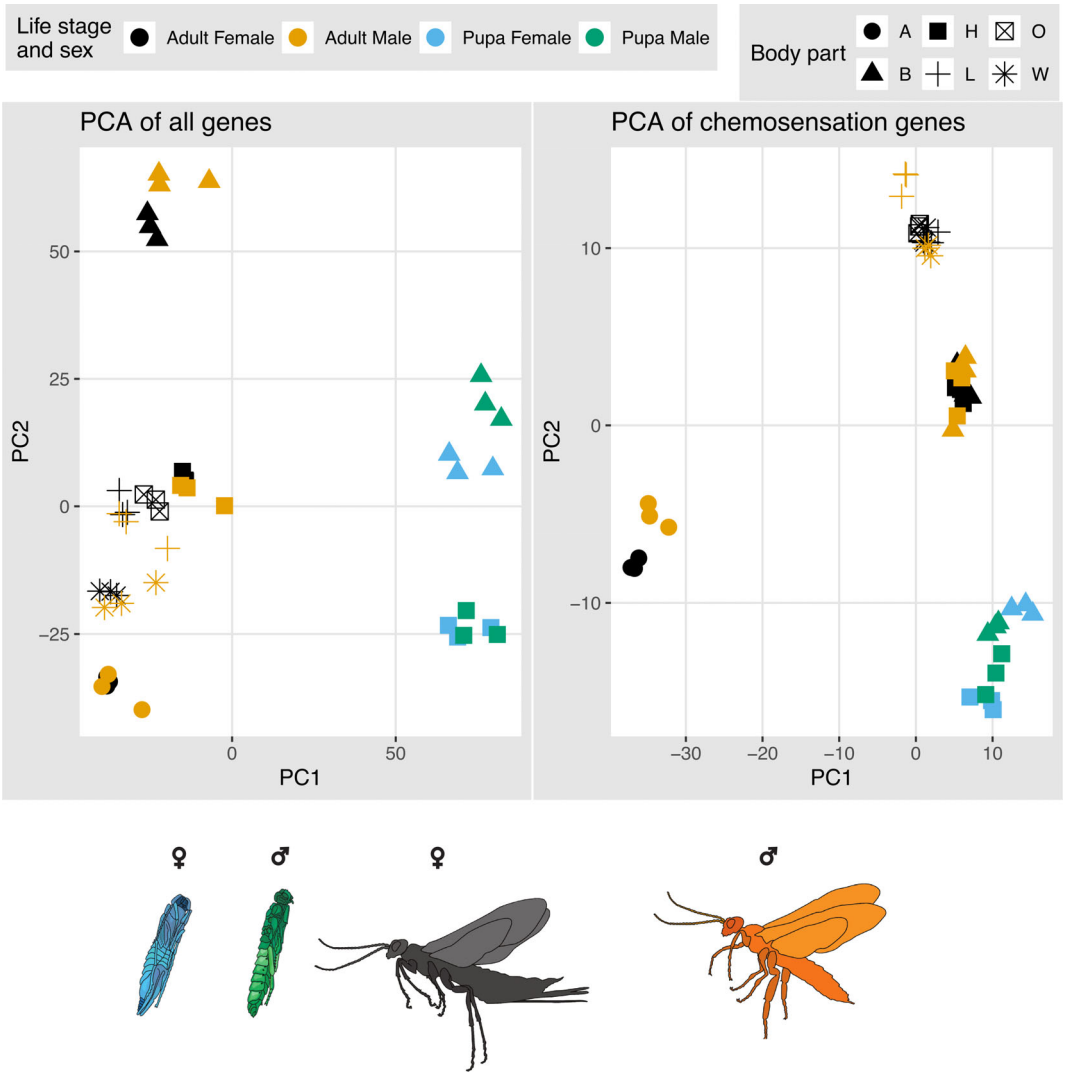

FIGURE 2 Principal component analysis (PCA) of all expressed genes (left) and chemosensation genes (right) of *Sirex noctilio*. Normalized count data were used with the plotPCA function of DESeq2 in R. Count data were produced from featureCounts and transformation of count data was accomplished through the built-in variance stabilizing transformation of DESeq2. Body parts: A, antennae (circle in key); B, body (triangle); H, head (square); L, leg (cross); O, ovipositor (cross in square); W, wings (star). Life stage and sex are as shown in the key and legend with insect figures below as follows: Adult female (black), adult male (orange), pupa female (blue) and pupae male (green).

included protein complex assembly, intracellular signalling, cell–cell adhesion, cell division, organelle transport, cell structure, signal transduction and ion channel regulation.

DISCUSSION

The first reported draft genome assembly of *S. noctilio* was produced in this study, providing data to identify and characterize the chemosensation genes of this economically significant pine tree pest. The genome assembly exhibited strong contiguity metrics with a high BUSCO score and high mapping rates of transcriptome sequences aligned to the assembled genome, indicating that the assembly is mostly complete. However, a high BUSCO completeness score and high transcriptome mapping rates might mask tandemly duplicated loci. To definitively confirm the completeness of such repetitive regions, further synteny or contiguity analyses should be conducted (Simão et al., 2015).

Fifteen *S. noctilio* transcriptomes were sequenced, and the expression of the identified chemosensation genes was investigated and compared between the different transcriptome datasets. Differentially expressed genes exhibited coherent functional clustering, and visualization of the full effect-size distributions indicated that the selection criteria applied were appropriately balanced to identify genes exhibiting high statistical significance and biologically meaningful effect sizes. However, it should be noted that independent validation of a subset of genes using qPCR would further have mitigated potential false positives and negatives in this study.

Global gene expression patterns were more distinct between developmental stages than across tissues. However, when focusing

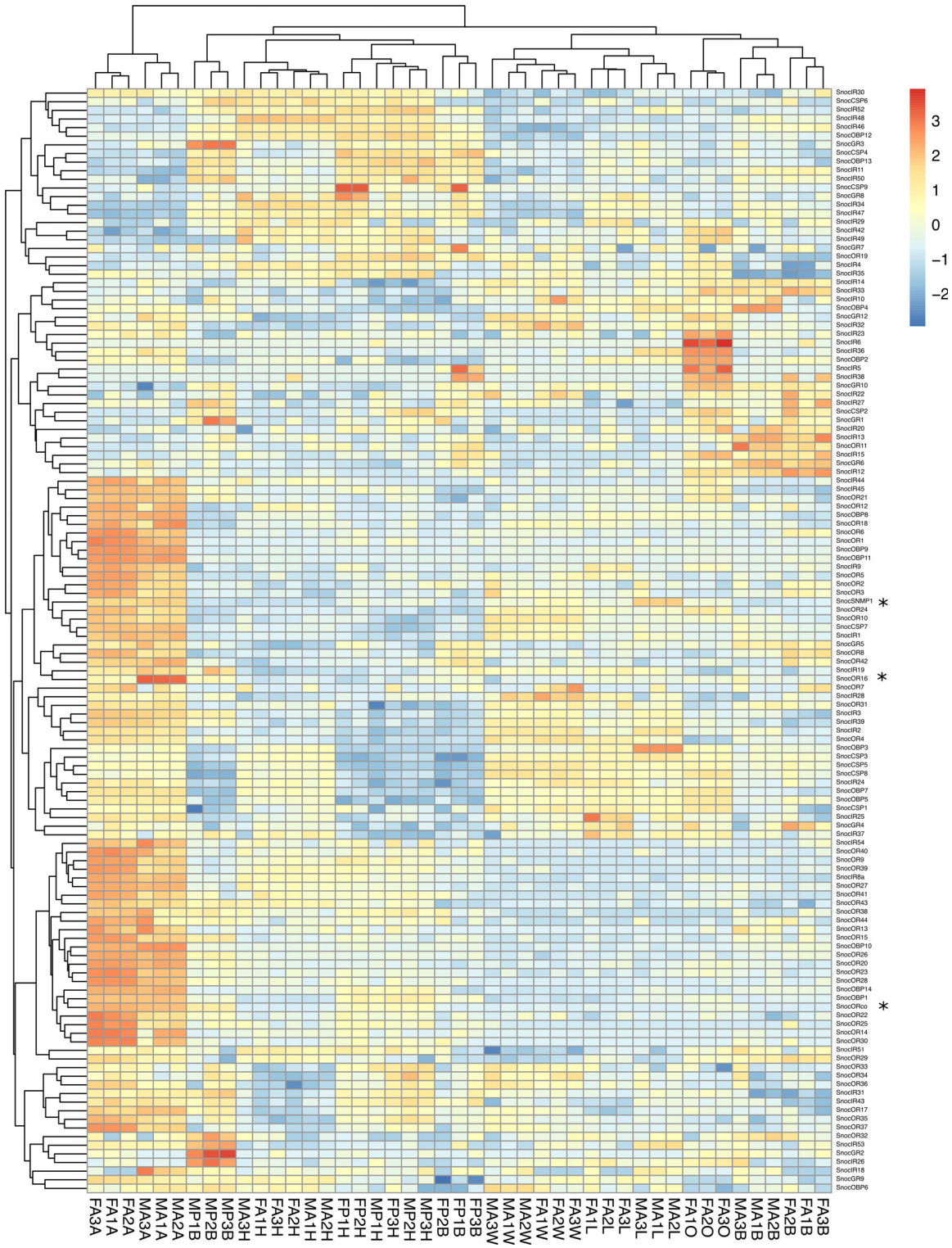

FIGURE 3 Heatmap of gene expression of chemosensation genes across all samples. Three replicates (indicated by the numerical values in each sample name) were sequenced for each sample. Normalized count data (as described for Figure 2) were used to generate the heatmap using the pheatmap function in R. Genes of interest which also display significant differential expression (log2FoldChange| ≥ 1, adj. $p < 0.05$) are annotated with star symbols (*). The effect of differential gene expression is particularly observed in the antennae (columns FA3A, FA1A, FA2A, MA3A, MA1A and MA2A), heads (MA3H, FA1H, FA3H, FA2H, MA1H, MA2H) and ovipositors (FA1O, FA2O, FA3O). These tissues are typically associated with chemosensation in insects.

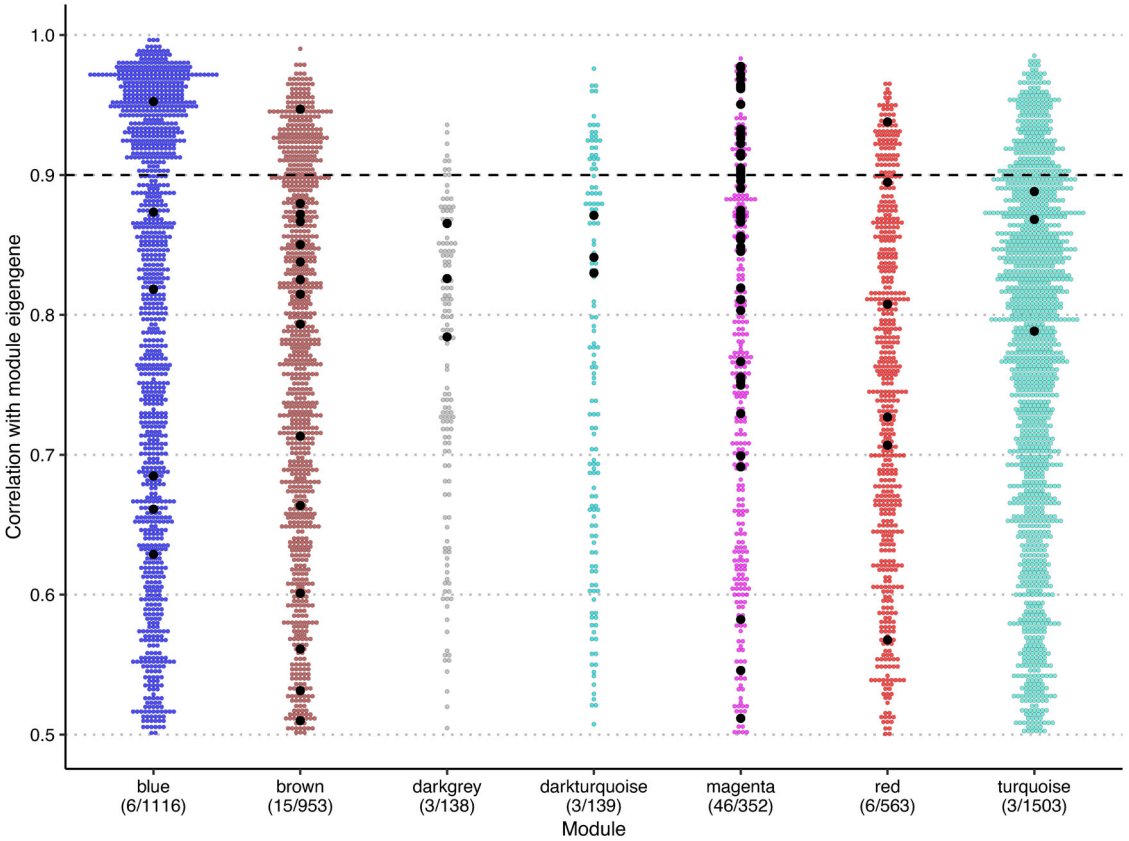

FIGURE 4 A hub-plot showing prominent modules and the hub genes within each module. The *y*-axis represents the correlation value between each gene and the eigengene of the module. The *x*-axis represents the modules containing the highest number of differentially expressed genes. The number of chemosensation genes in relation to the total number of genes per module is shown in brackets under each module name. Black dots represent chemosensation genes within each module. A dashed line separates genes based on the correlation value of 0.9 across all modules.

on chemosensation genes, antennae formed a distinct cluster separate from all other tissues and developmental stages. This underscores the key role of antennae in chemosensation, particularly olfaction. Compared to other tissues, antennae exhibited the greatest number of differentially expressed chemosensory genes. Of the six gene families studied, ORs showed the greatest expression bias in antennal tissues. Per-gene differential expression patterns are further discussed below.

WGCNA enabled the identification of co-expression modules in *S. noctilio*. The magenta module contained the greatest amount of chemosensory hub genes which are likely co-expressed and co-regulated. GSVA provided additional evidence that the antennae exhibited enriched gene expression patterns compared to other tissues studied. Non-chemosensory genes identified in the magenta module were predicted to function as transcription factors, which might be important targets for further investigation of the regulation of chemosensation in *S. noctilio*. Other key functions related to the chemosensation process were also reflected in the predicted functions of these non-chemosensory co-expressed genes. These included protein signal transduction and ion channel regulation. Ion channels are vital for operation of the chemosensory system (Wicher, 2012). Some chemoreceptors function as ion channels while others may trigger intracellular signalling cascades which lead to the activation of ion channels.

The co-receptor, SnocOrco, was expressed in all tissues, with expression levels surpassing those of other chemoreceptor genes. The antenna-biased expression of ORs, along with the high expression levels of ORco, aligns with findings from other studies (Boronat-Garcia et al., 2022; Gu et al., 2014; Jiang et al., 2021; Robertson, Waterhouse, et al., 2018). ORco is an essential co-factor for all other ORs, enabling the precise detection of chemical stimuli, which may be critical for a range of ecologically important chemosensory functions, such as oviposition-site selection and mate location (Fleischer et al., 2018).

SnocOR16 was expressed at significantly higher levels in male antennae, a result also observed by Guo et al. (2021). Phylogenetically, this receptor clustered at the base of the 9-exon subfamily, which is thought to be involved in the reception of cuticular hydrocarbons (CHCs) and has been implicated in social interactions within Hymenoptera (Karpe et al., 2016; Obiero et al., 2021; Würf et al., 2020; Yan et al., 2020; Zhou et al., 2012). CHCs, which cover the insect cuticle, serve to protect against desiccation and mediate social behaviours, ranging from courtship and aggregation to colony organization and conspecific communication, especially in eusocial insects (Blomquist & Bagnères, 2010; Ferguson et al., 2020; Sharma et al., 2015; Yan et al., 2020; Zhou et al., 2012; Zhou et al., 2015).

Ants, for example, can discriminate between CHCs without direct contact, suggesting that CHCs may function as potential olfactory cues detectable by ORs in antennae (Pask et al., 2017). In *S. noctilio*, females produce monoalkenes, a specific group of desaturated CHCs, on their cuticles that induce male copulation (Böröczky et al., 2009) and may also play a role in other recently reported mating behaviours. For example, Queffelec et al. (2021) found that female attractiveness decreases with age and that male *S. noctilio* wasps exhibit potential mate choice. They proposed that CHC production by female wasps could decline with age, making older females less attractive and reducing their mating opportunities. The phylogenetic position of SnocOR16 might indicate its involvement in reception of CHCs involved in mating and reproduction. SnocOR16 was also identified as a hub gene in one of the most prominent modules identified by WGCNA. Hub genes represent central nodes within a gene co-expression network and are genes that are highly connected to many other genes within the same co-expression module. These genes are potentially key drivers or regulators of specific processes or traits being studied. Similar studies have been performed to identify genes playing key regulatory roles in the chemosensory process of *Glossina morsitans morsitans* (Gakii et al., 2021). Deorphanization studies involving SnocOR16 and CHC ligands could shed light on the potential role of this receptor in the sexual behavioural ecology of *S. noctilio*.

The results of phylogenetic and gene expression analyses suggest that additional ORs may be involved in sex-specific behaviours in *S. noctilio*. SnocOR6, 9, 22, 23, 24, 28 and 39 were significantly upregulated in female antennae compared to male antennae. Four of these seven genes belong to a group of 13 *S. noctilio* ORs (SnocOR9, 13, 14, 23, 25, 28, 29, 30, 38, 39, 40, 41 and 44), which may result from species-specific expansion and form a unique phylogenetic clade within the OR gene family. Six of these *S. noctilio* genes (SnocOR9, 14, 28, 29, 31 and 40) are orthologous to ORs from the European spruce bark beetle, *Ips typographus* (Hou et al., 2021). It is predicted that the receptors encoded by these bark beetle genes interact with terpenoids from host conifer trees and symbiotic fungi (Hou et al., 2021). This finding is potentially significant, as the *S. noctilio* orthologues may serve a similar function, responding to volatiles of related host trees and symbiotic fungi. Whether this is also the case in *S. noctilio* requires further experimental studies. Females are expected to express these receptors at higher levels than males, as they must locate suitable hosts for oviposition. SnocOR9 and SnocOR14 were also hub genes in the magenta module identified via WGCNA, which provides more evidence of the significance and co-expression of these receptors.

SnocOR24 was significantly upregulated in female antennae, with biased expression also observed in ovipositor tissues. Phylogenetically, SnocOR24 is positioned at the base of a large clade that includes the honeybee (*Apis mellifera*) gene, AmelOR11. AmelOR11 is the honeybee OR responsible for detecting the queen pheromone component, 9-oxo-2-decenoic acid, and is one of the few olfactory receptors that have been deorphanized (Liu et al., 2018; Wanner et al., 2007). The orthology to the honeybee pheromone receptor, combined with its female-biased expression patterns, suggests that SnocOR24 may play a role in *S. noctilio* female sexual behaviour.

SnocOR20 and SnocOR21 exhibited high expression levels in antennae, with SnocOR20 showing significant upregulation in antennae compared to other tissues. These receptors were also found to be hub genes in the magenta module resulting from WGCNA, which means they are co-expressed with other significantly differentially expressed OR genes such as SnocOR9, SnocOR14 and SnocOR16. SnocOR20 and SnocOR21 are orthologues of the *A. mellifera* gene, AmelOR151, which encodes a receptor that responds to the floral terpene, linalool (Claudianos et al., 2014). Linalool has several reported biological functions, including antimicrobial, antioxidant, larvicidal, insecticidal, pollinator attractant and repellent activities (Fujiwara et al., 2017; Kamatou & Viljoen, 2008). The symbiotic fungus *A. areolatum*, associated with *S. noctilio*, produces a four-compound blend that includes linalool, to which female wasps are attracted (Faal et al., 2021). In the wheat stem sawfly, *C. cinctus*, behavioural studies have shown that host selection by ovipositing females is influenced by host volatiles such as (Z)-3-hexenyl acetate and linalool (Achhami et al., 2021). Achhami et al. (2021) suggested that high levels of linalool in the host plant may deter oviposition by *C. cinctus* females. Therefore, SnocOR20 and SnocOR21 may play a role in recognizing ecologically important symbiont or host volatiles.

Normalized gene count data showed that male leg tissues clustered separately from wings, ovipositor and female leg tissues. Adult *S. noctilio* exhibit sexual dimorphism, with males possessing larger hind legs than females. Hind legs of *S. noctilio* have been hypothesized to be the site of synthesis and/or storage of a putative male-produced pheromone blend that may function as an aggregation pheromone (Guignard et al., 2021). This pheromone could play a role in lek formation among male *S. noctilio*. Therefore, genes such as SnocOBP3 and SnocSNMP1, which show enhanced expression in male leg tissues, are of particular interest as target genes for potential future control strategies.

The expression of SnocOBP3 in adult male legs suggests its potential role in male-specific behaviours. Previous transcriptome studies have highlighted that OBPs with leg-biased expression might mediate host-seeking behaviours (Ma et al., 2018; Song et al., 2018; Sun et al., 2017), although, unlike SnocOBP3, the OBPs identified in those studies were not sex-biased. The male leg-biased expression of SnocOBP3 could imply a function in pheromone detection and/or the release and storing of chemical compounds (Pelosi et al., 2017). A relevant comparison can be made with *Bactrocera dorsalis* (oriental fruit fly) females, which produce 4-allyl-2,6-dimethoxyphenol (4-DMP), a phenolic derivative that acts as a close-range sex pheromone attractive to males. In *B. dorsalis*, male antennae do not respond to 4-DMP during electroantennography (EAG) (Shen et al., 2019), but an electrophysiological response is observed in male mid-legs (Hu et al., 2021). Since male *B. dorsalis* legs are in close contact with females during copulation, the interaction of 4-DMP with male legs likely plays a critical role in mating behaviour. In fact, exposure to 4-DMP led to downregulation of BdorOBP2, an OBP strongly binding to 4-DMP, suggesting it is a transporter for this pheromone (Hu et al., 2021). The reduced transcript levels of BdorOBP2 corresponded to a decreased behavioural response in male fruit flies, which could represent a

mechanism for preventing overstimulation of the chemosensory system (von der Weid et al., 2015). Given these parallels, it would be valuable to investigate the possible involvement of SnocOBP3 in the perception of CHCs in male *S. noctilio* leg tissues during copulation.

The male leg-biased expression of SnocSNMP1 along with its antennal expression aligns with the hypothesis that SNMP1 is involved in pheromone detection, working in conjunction with ORs to recognize lipophilic pheromones (Forstner et al., 2008). The expression pattern is similar to that observed in the Asian honeybee, *Apis cerana cerana*, where SNMP1 is highly expressed in both antennae and legs (Hu et al., 2013). In the case of *A. cerana cerana*, the presence of SNMP1 in legs was linked to its involvement in various sensory processes. The identification of male *S. noctilio* hind legs as a potential site for the synthesis and/or storage of a putative aggregation pheromone (Guignard et al., 2021) further supports the possibility that SnocSNMP1 expression in male legs plays a role in pheromone detection. Taken together, the phylogenetic and gene expression data suggest that SnocSNMP1 may be implicated in the sex-specific behaviours of *S. noctilio*, potentially contributing to mating and aggregation processes.

CONCLUSIONS

Phylogenetic analyses that include species of early diverged lineages, such as *S. noctilio*, contribute particularly to our understanding of the evolutionary histories of multigene families in Hymenoptera. Furthermore, the specialization of chemosensory system genes to particular habitats and ecological niches warrants further interest, especially for invasive and pest species. These genes encoding chemosensory proteins are potential targets for insect population control strategies. The phylogenetic relationships and expression patterns of chemosensation genes identified in this study revealed several promising targets for further investigation into the ecological specialization and success of *S. noctilio*. Specifically, the study found that SnocOR16 and SnocSNMP1 exhibit tissue- and sex-biased expression patterns. This research advances our understanding of the peri-receptive processes involved in *S. noctilio* chemosensation and lays the groundwork for future chemical ecology studies and functional deorphanization of the chemosensation genes of this invasive species.

AUTHOR CONTRIBUTIONS

Alisa Postma: Conceptualization; investigation; writing – original draft; methodology; validation; visualization; writing – review and editing; software; formal analysis; project administration; data curation. **Leandri Klynsmith:** Conceptualization; investigation; writing – original draft; methodology; validation; visualization; writing – review and editing; formal analysis; data curation. **Tuan A. Duong:** Conceptualization; methodology; writing – review and editing; supervision. **Jeremy Dean Allison:** Conceptualization; methodology; writing – review and editing; supervision. **Werner Smidt:** Methodology; software; writing – review and editing; visualization. **Robert M. Waterhouse:** Conceptualization; methodology; software; writing – review and editing. **Peter Lesny:** Investigation; resources; writing – review and editing. **Jan Philip Oeyen:** Investigation; resources; writing – review and editing. **Malte Petersen:** Investigation; writing – review and editing; resources. **Sebastian Martin:** Investigation; resources. **Shanlin Liu:** Investigation; writing – review and editing; resources. **Xin Zhou:** Investigation; resources. **Tanja Ziesmann:** Investigation. **Alexander Donath:** Investigation; writing – review and editing; resources. **Christoph Mayer:** Investigation; writing – review and editing; resources. **Bernhard Misof:** Investigation; writing – review and editing; resources. **Oliver Niehuis:** Investigation; writing – review and editing; resources. **Ralph S. Peters:** Investigation; writing – review and editing; resources. **Lars Podsiadlowski:** Investigation; resources. **Martin P. A. Coetzee:** Investigation; writing – review and editing. **Fourie Joubert:** Resources; methodology; writing – review and editing; supervision. **Bernard Slippers:** Conceptualization; resources; writing – review and editing; project administration; supervision; funding acquisition.

ACKNOWLEDGEMENTS

We thank Jeanne Wilbrandt (ORCID: 0000-0002-0363-3837, Leibniz FLI Jena, Core Facility Life Science Computing) for help managing the genome sequence data. We thank Melissa Fierke (ORCID: 0000-0002-1561-0720) and Christopher Foelker (State University of New York, College of Environmental Sciences and Forestry) for providing *S. noctilio* specimens. Sincere thanks to Glenda Brits who assisted us with graphic design for the figures in this manuscript. We thank a member of the Tree Protection Cooperative Program for logs supplied to obtain *S. noctilio* adults, larvae and pupae.

FUNDING INFORMATION

This study was supported by the DSTI NRF Centre of Excellence in Plant Health Biotechnology and the Tree Protection Cooperative Programme in South Africa as well as the Leibniz Graduate School on Genomic Biodiversity Research and the German Research Foundation (DFG) (MI 649/16–1, NI-1387/3–1). R.M.W. acknowledges support from the Swiss National Science Foundation grants 186397 and 202669.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Raw RNA-sequencing data are available at the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database under BioProject ID PRJNA982519: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA982519/ (Postma et al., 2025a). Genome assembly, annotation and predicted protein sequences: https://doi.org/10.5061/dryad.tdz08kqbp (Postma et al., 2025b). All code to reproduce the results and figures is available at https://github.com/AlisaPostma/SirexChemosensation (Postma et al., 2025c).

ORCID

Alisa Postma https://orcid.org/0000-0002-7842-8096
Leandri Klynsmith https://orcid.org/0000-0001-9428-8524
Robert M. Waterhouse https://orcid.org/0000-0003-4199-9052
Peter Lesny https://orcid.org/0000-0001-7601-9873
Jan Philip Oeyen https://orcid.org/0000-0002-0116-2234
Malte Petersen https://orcid.org/0009-0000-2173-989X
Sebastian Martin https://orcid.org/0000-0003-3171-4420
Shanlin Liu https://orcid.org/0000-0001-8118-8313
Alexander Donath https://orcid.org/0000-0001-5618-0547
Christoph Mayer https://orcid.org/0000-0001-5104-6621
Bernhard Misof https://orcid.org/0000-0003-4175-6798
Oliver Niehuis https://orcid.org/0000-0003-4253-1849
Ralph S. Peters https://orcid.org/0000-0001-7784-9203
Lars Podsiadlowski https://orcid.org/0000-0001-7786-8930
Bernard Slippers https://orcid.org/0000-0003-1491-3858

REFERENCES

Achhami, B.B., Reddy, G.V.P., Hofland, M.L., Sherman, J.D., Peterson, R.K.D. & Weaver, D.K. (2021) Plant volatiles and oviposition behavior in the selection of barley cultivars by wheat stem sawfly (Hymenoptera: Cephidae). *Environmental Entomology*, 50, 940–947.

Allison, J.D. (2021) Chemical ecology and trapping technique for *Sirex noctilio*. In: Hajek, A.E., Haavik, L.J. & Stephen, F.M. (Eds.) *Biology and ecology of Sirex noctilio in North America*. Morgantown: USDA Forest Service. 22–29.

Andrews, S. (2010) *FastQC: a quality control tool for high throughput sequence data*. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Blomquist, G.J. & Bagnères, A.-G. (2010) *Insect hydrocarbons: biology, biochemistry, and chemical ecology*. Cambridge: Cambridge University Press.

Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D. & Pirovano, W. (2011) Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*, 27, 578–579.

Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120.

Böröczky, K., Crook, D.J., Jones, T.H., Kenny, J.C., Zylstra, K.E., Mastro, V.C. et al. (2009) Monoalkenes as contact sex pheromone components of the Woodwasp *Sirex noctilio*. *Journal of Chemical Ecology*, 35, 1202–1211.

Boronat-Garcia, A., Iben, J., Dominguez-Martin, E. & Stopfer, M. (2022) Identification and analysis of odorant receptors expressed in the two main olfactory organs, antennae and palps, of *Schistocerca americana*. *Scientific Reports*, 12, 22628.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. et al. (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, 10, 421.

Campbell, M.S., Holt, C., Moore, B. & Yandell, M. (2014) Genome annotation and curation using MAKER and MAKER-P. *Current Protocols in Bioinformatics*, 48, 4.11.1–4.11.39.

Capella-Gutiérrez, S., Silla-Martínez, J.M. & Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15), 1972–1973.

Claudianos, C., Lim, J., Young, M., Yan, S., Cristino, A.S., Newcomb, R.D. et al. (2014) Odor memories regulate olfactory receptor expression in the sensory periphery. *European Journal of Neuroscience*, 39(10), 1642–1654.

Emms, D.M. & Kelly, S. (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, 20, 238.

Ewels, P., Magnusson, M., Lundin, S. & Käller, M. (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19), 3047–3048.

Eyun, S.-I., Soh, H.Y., Posavi, M., Munro, J.B., Hughes, D.S.T., Murali, S.C. et al. (2017) Evolutionary history of chemosensory-related gene families across the Arthropoda. *Molecular Biology and Evolution*, 34, 1838–1862.

Faal, H., Cha, D., Hajek, A. & Teale, S. (2021) A double-edged sword: *Amylostereum areolatum* odors attract both *Sirex noctilio* (Hymenoptera: Siricidae) and its parasitoid, *Ibalia leucospoides*. *Fungal Ecology*, 54, 101108.

Ferguson, S.T., Park, K.Y., Ruff, A.A., Bakis, I. & Zwiebel, L.A.-O. (2020) Odor coding of nestmate recognition in the eusocial ant *Camponotus floridanus*. *The Journal of Experimental Biology*, 223(2), jeb.215400.

Fernández Ajó, A.A., Martínez, A.S., Villacide, J.M. & Corley, J.C. (2015) Behavioural response of the woodwasp *Sirex noctilio* to volatile emissions of its fungal symbiont. *Journal of Applied Entomology*, 139, 654–659.

Fleischer, J., Pregitzer, P., Breer, H. & Krieger, J. (2018) Access to the odor world: olfactory receptors and their role for signal transduction in insects. *Cellular and Molecular Life Sciences*, 75, 485–508.

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

Fujiwara, G.M., Annies, V., de Oliveira, C.F., Lara, R.A., Gabriel, M.M., Betim, F.C. et al. (2017) Evaluation of larvicidal activity and ecotoxicity of linalool, methyl cinnamate and methyl cinnamate/linalool in combination against *Aedes aegypti*. *Ecotoxicology and Environmental Safety*, 139, 238–244.

Gakii, C., Bwana, B.K., Mugambi, G.G., Mukoya, E., Mireji, P.O. & Rimiru, R. (2021) In silico-driven analysis of the *Glossina morsitans* morsitans antennae transcriptome in response to repellent or attractant compounds. *PeerJ*, 9, e11691.

Gu, S.-H., Sun, L., Yang, R.-N., Wu, K.-M., Guo, Y.-Y., Li, X.-C. et al. (2014) Molecular characterization and differential expression of olfactory genes in the antennae of the black cutworm moth *Agrotis ipsilon*. *PLoS One*, 9, e103420.

Guignard, Q., Bouwer, M., Slippers, B. & Allison, J. (2021) Biology of a putative male aggregation-sex pheromone in *Sirex noctilio* (Hymenoptera: Siricidae). *PLoS One*, 15, e0244943.

Guo, B., Hao, E., Qiao, H., Wang, J., Wu, W., Zhou, J. et al. (2021) Antennal transcriptome analysis of olfactory genes and characterizations of odorant binding proteins in two woodwasps, *Sirex noctilio* and *Sirex nitobei* (Hymenoptera: Siricidae). *BMC Genomics*, 22, 172.

Hänzelmann, S., Castelo, R. & Guinney, J. (2013) GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*, 14, 7.

Hayes, R.A., Griffiths, M.W. & Nahrung, H.F. (2015) Electrophysiological activity of the *Sirex noctilio* ovipositor: you know the drill? *Journal of Asia-Pacific Entomology*, 18, 165–168.

Hoff, K.J., Lange, S., Lomsadze, A., Borodovsky, M. & Stanke, M. (2016) BRAKER1: unsupervised RNA-seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*, 32(5), 767–769.

Hou, X.-Q., Yuvaraj, J.K., Roberts, R.E., Zhang, D.-D., Unelius, C.R., Löfstedt, C. et al. (2021) Functional evolution of a bark beetle odorant receptor clade detecting monoterpenoids of different ecological origins. *Molecular Biology and Evolution*, 38, 4934–4947.

Hu, L., Chen, B., Liu, K., Yu, G., Chen, Y., Dai, J. et al. (2021) OBP2 in the Midlegs of the male *Bactrocera dorsalis* is involved in the perception of the female-biased sex pheromone 4-Allyl-2,6-dimethoxyphenol. *Journal of Agricultural and Food Chemistry*, 69, 126–134.

Hu, Y.-Y., Xu, S.-F., Wubie, A., Li, W., Huang, J., Guo, Z.-B. et al. (2013) Molecular characterization and tissue localization of sensory neuron membrane protein from Chinese honey bee, *Apis cerana cerana* (Hymenoptera: Apidae). *Applied Entomology and Zoology*, 48, 533–545.

Hurley, B.P., Slippers, B. & Wingfield, M.J. (2007) A comparison of control results for the alien invasive woodwasp, *Sirex noctilio*, in the southern hemisphere. *Agricultural and Forest Entomology*, 9, 159–171.

Jiang, X.-C., Liu, S., Jiang, X.-Y., Wang, Z.-W., Xiao, J.-J., Gao, Q. et al. (2021) Identification of olfactory genes from the greater wax moth by antennal transcriptome analysis. *Frontiers in Physiology*, 12, 663040.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., Mcanulla, C. et al. (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 30, 1236–1240.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A. & Jermiin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14, 587–589.

Kamatou, G.P. & Viljoen, A. (2008) Linalool – a review of a biologically active compound of commercial importance. *Natural Product Communications*, 3, 1183–1192.

Karpe, S.D., Jain, R., Brockmann, A. & Sowdhamini, R. (2016) Identification of complete repertoire of *Apis florea* odorant receptors reveals complex orthologous relationships with *Apis mellifera*. *Genome Biology and Evolution*, 8(9), 2879–2895.

Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780.

Kim, D., Langmead, B. & Salzberg, S.L. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, 12, 357–360.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. & Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, 14, R36.

Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E.L.L. (2001) Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology*, 305, 567–580.

Langfelder, P. & Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 9, 559.

Langfelder, P. & Horvath, S. (2012) Fast R functions for robust correlations and hierarchical clustering. *Journal of Statistical Software*, 46, 1–17.

Lee, E., Helt, G.A., Reese, J.T., Munoz-Torres, M.C., Childers, C.P., Buels, R.M. et al. (2013) Web Apollo: a web-based genomic annotation editing platform. *Genome Biology*, 14, R93.

Liao, Y., Smyth, G.K., Shi, W. & Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923–930.

Liu, J.F., Yang, L., Li, M., Xujiang, H., Wang, Z. & Zeng, Z.-J. (2018) Cloning and expression pattern of odorant receptor 11 in Asian honeybee drones, *Apis cerana* (Hymenoptera, Apidae). *Journal of Asia-Pacific Entomology*, 22, 110–116.

Lizana, P., Mutis, A., Quiroz, A. & Venthur, H. (2022) Insights into chemosensory proteins from non-model insects: advances and perspectives in the context of pest management. *Frontiers in Physiology*, 13, 924750.

Love, M.I., Huber, W. & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15, 550.

Ma, L., Li, Z., Zhang, W., Cai, X., Luo, Z., Zhang, Y. et al. (2018) The odorant binding protein 6 expressed in sensilla Chaetica displays preferential binding affinity to host plants volatiles in *Ectropis obliqua*. *Frontiers in Physiology*, 9, 534.

Nguyen, L.-T., Schmidt, H.A., von Haeseler, A. & Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32, 268–274.

Obiero, G.F., Pauli, T., Geuverink, E., Veenendaal, R., Niehuis, O. & Gross-Wilde, E. (2021) Chemoreceptor diversity in Apoid wasps and its reduction during the evolution of the pollen-collecting lifestyle of bees (Hymenoptera: Apoidea). *Genome Biology and Evolution*, 13(3), evaa269.

Oeyen, J.P., Baa-Puyoulet, P., Benoit, J.B., Beukeboom, L.W., Bornberg-Bauer, E., Buttstedt, A. et al. (2020) Sawfly genomes reveal evolutionary acquisitions that fostered the mega-radiation of parasitoid and eusocial Hymenoptera. *Genome Biology and Evolution*, 12(7), 1099–1188.

Pask, G.M., Slone, J.D., Millar, J.G., Das, P., Moreira, J.A., Zhou, X. et al. (2017) Specialized odorant receptors in social insects that detect cuticular hydrocarbon cues and candidate pheromones. *Nature Communications*, 8, 297.

Pelosi, P., Iovinella, I., Zhu, J., Wang, G. & Dani, F.R. (2017) Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. *Biological Reviews*, 93, 184–200.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.-C., Mendell, J.T. & Salzberg, S.L. (2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, 33, 290–295.

Peters, R.S., Krogmann, L., Mayer, C., Donath, A., Gunkel, S., Meusemann, K. et al. (2017) Evolutionary history of the Hymenoptera. *Current Biology*, 27, 1013–1018.

Petersen, T.N., Brunak, S., von Heijne, G. & Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, 8, 785.

Postma, A., Klynsmith, L., Duong, T.A., Allison, J.D., Smidt, W., Waterhouse, R.M. et al. (2025a) *Genome and transcriptome-based identification and expression profiling of chemosensory gene families across developmental stages and tissues in Sirex noctilio (Hymenoptera: Siricidae)*. Sequence Read Archive (SRA). Available at: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA982519

Postma, A., Klynsmith, L., Duong, T.A., Allison, J.D., Smidt, W., Waterhouse, R.M. et al. (2025b) *Genome and transcriptome-based identification and expression profiling of chemosensory gene families across developmental stages and tissues in Sirex noctilio (Hymenoptera: Siricidae)*. Dryad Digital Repository. Available from: https://doi.org/10.5061/dryad.tdz08kqbp

Postma, A., Klynsmith, L., Duong, T.A., Allison, J.D., Smidt, W., Waterhouse, R.M. et al. (2025c) *Genome and transcriptome-based identification and expression profiling of chemosensory gene families across developmental stages and tissues in Sirex noctilio (Hymenoptera: Siricidae)*. GitHub. Available at: https://github.com/AlisaPostma/SirexChemosensation

Queffelec, J., Allison, J.D., Slippers, B. & Greeff, J.M. (2021) *Female and male Sirex noctilio use age and size to select a mate*. bioRxiv, 2021.02.14.431120.

Robertson, H.M., Baits, R.L., Walden, K.K.O., Wada-Katsumata, A. & Schal, C. (2018) Enormous expansion of the chemosensory gene repertoire in the omnivorous German cockroach *Blattella germanica*. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 330, 265–278.

Robertson, H.M., Waterhouse, R.M., Walden, K.K.O., Ruzzante, L., Reijnders, M.J.M.F., Coates, B.S. et al. (2018) Genome sequence of the wheat stem sawfly, *Cephus cinctus*, a primitive hymenopteran and wheat pest, illuminates evolution of hymenopteran chemoreceptors. *Genome Biology and Evolution*, 10(11), 2997–3011.

Sharma, K.R., Enzmann, B.L., Schmidt, Y., Moore, D., Jones, G.R., Parker, J. et al. (2015) Cuticular hydrocarbon pheromones for social behavior and their coding in the ant antenna. *Cell Reports*, 12(8), 1261–1271.

Shen, J., Hu, L., Zhou, X., Dai, J., Chen, B. & Li, S. (2019) Allyl-2,-6-dimethoxyphenol, a female-biased compound, is robustly attractive to conspecific males of *Bactrocera dorsalis* at close range. *Entomologia Experimentalis et Applicata*, 167, 811–819.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V. & Zdobnov, E.M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31, 3210–3212.

Slippers, B., de Groot, P. & Wingfield, M. (2012) *The Sirex woodwasp and its fungal symbiont: research and management of a worldwide invasive pest*. Dordrecht: Springer.

Song, Y.-Q., Sun, H.-Z. & Du, J. (2018) Identification and tissue distribution of chemosensory protein and odorant binding protein genes in *Tropidothorax elegans* distant (Hemiptera: Lygaeidae). *Scientific Reports*, 8, 7803.

Stanke, M., Diekhans, M., Baertsch, R. & Haussler, D. (2008) Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics*, 24, 637–644.

Sun, L., Wang, Q., Wang, Q., Dong, K., Xiao, Y. & Zhang, Y.-J. (2017) Identification and characterization of odorant binding proteins in the forelegs of *Adelphocoris lineolatus* (Goeze). *Frontiers in Physiology*, 8, 735.

von der Weid, B., Rossier, D., Lindup, M., Tuberosa, J., Widmer, A., Col, J.D. et al. (2015) Large-scale transcriptional profiling of chemosensory neurons identifies receptor-ligand pairs in vivo. *Nature Neuroscience*, 18(10), 1455–1463.

Wanner, K.W., Nichols, A.S., Walden, K.K.O., Brockmann, A., Luetje, C.W. & Robertson, H.M. (2007) A honey bee odorant receptor for the queen substance 9-oxo-2-decenoic acid. *Proceedings of the National Academy of Sciences*, 104, 14383–14388.

Wicher, D. (2012) Functional and evolutionary aspects of chemoreceptors. *Frontiers in Cellular Neuroscience*, 6, 48.

Würf, J., Pokorny, T., Wittbrodt, J., Millar, J.G. & Ruther, J. (2020) Cuticular hydrocarbons as contact sex pheromone in the parasitoid wasp *Urolepis rufipes*. *Frontiers in Ecology and Evolution*, 8, 180.

Yan, H., Jafari, S., Pask, G., Zhou, X., Reinberg, D. & Desplan, C. (2020) Evolution, developmental expression and function of odorant receptors in insects. *The Journal of Experimental Biology*, 223(suppl 1), jeb.208215.

Yu, G., Smith, D.K., Zhu, H., Guan, Y. & Lam, T.T.-Y. (2016) Ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, 8, 28–36.

Zerbino, D.R. & Birney, E. (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, 18, 821–829.

Zhou, X., Rokas, A., Berger, S.L., Liebig, J., Ray, A. & Zwiebel, L.J. (2015) Chemoreceptor evolution in Hymenoptera and its implications for the evolution of eusociality. *Genome Biology and Evolution*, 7, 2407–2416.

Zhou, X., Slone, J.D., Rokas, A., Berger, S.L., Liebig, J., Ray, A. et al. (2012) Phylogenetic and transcriptomic analysis of chemosensory receptors in a pair of divergent ant species reveals sex-specific signatures of odor coding. *PLoS Genetics*, 8, e1002930.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Heatmap depicting differential expression results for OR genes in *Sirex noctilio*.

Figure S2. Heatmap depicting differential expression results for GR genes in *Sirex noctilio*.

Figure S3. Phylogenetic tree of GR genes from *Sirex noctilio* (Snoc), *Anopheles gambiae* (Agam), *Apis mellifera* (Amel), *Athalia rosae* (Aros), *Bombyx mori* (Bmor), *Bombus terristris* (Bter), *Cephus cinctus* (Ccin), *Drosophila melanogaster* (Dmel), *Nasonia vitripennis* (Nvit), *Orussus abietinus* (Oabi), *Pogonomyrmex barbatus* (Pbar) and *Tribolium castaneum* (Tcas). Sub-families of interest are highlighted and include: pheromone receptors, sugar and fructose receptors and carbon dioxide receptors.

Figure S4. Heatmap depicting differential expression results for IR genes in *Sirex noctilio*.

Figure S5. Phylogenetic tree of IR genes from *Sirex noctilio* (Snoc), *Apis mellifera* (Amel), *Athalia rosae* (Aros), *Bombus terristris* (Bter), *Cephus cinctus* (Ccin), *Drosophila melanogaster* (Dmel), *Nasonia vitripennis* (Nvit), *Orussus abietinus* (Oabi) and *Pogonomyrmex barbatus* (Pbar). Sub-families of interest are highlighted and include: Divergent receptors and Co-receptors (Ir8a and Ir25a).

Figure S6. Heatmap depicting differential expression results for OBP genes in *Sirex noctilio*.

Figure S7. Phylogenetic tree of OBP genes from *Sirex noctilio* (Snoc), *Apis mellifera* (Amel), *Cephus cinctus* (Ccin), *Nasonia vitripennis* (Nvit) and *Solenopsis invicta* (Sinv). Sub-families of interest are highlighted and include a bee-specific clade, a wasp-specific clade and an ant-specific clade.

Figure S8. GSVA enrichment of the Magenta module.

Figure S9. Heatmap of GSVA enrichment of the Magenta module.

Figure S10. Volcano plot.

Table S1. (a) RNA sample sequencing quality metrics. (b) RNA sequencing library information.

Table S2. DNA sample and genome sequencing information.

Table S3. Chemosensation gene information.

Table S4. Information on sequenced Symphyta genomes.

Table S5. Numbers of chemosensory genes in different insect species.

Table S6. Mapping rate of RNA libraries to the *Sirex* genome.

Table S7. Genes in magenta module.

File S1. DNA extraction phenol chloroform protocol.

File S2. Total RNA extraction trizol.

File S3. Surface sterilization and sterilising wash procedures for RNA extraction from insects.

File S4. Gut removal from insects.

How to cite this article: Postma, A., Klynsmith, L., Duong, T.A., Allison, J.D., Smidt, W., Waterhouse, R.M. et al. (2026) Genome and transcriptome-based identification and expression profiling of chemosensory gene families across developmental stages and tissues in *Sirex noctilio* (Hymenoptera: Siricidae). *Insect Molecular Biology*, 1–14. Available from: https://doi.org/10.1111/imb.70029