frontiers | Frontiers in Ecology and Evolution

TYPE Original Research
PUBLISHED 12 January 2023
DOI 10.3389/fevo.2022.1065947

OPEN ACCESS

EDITED BY
Alberto Maria Cattaneo,
Swedish University of Agricultural Sciences, Sweden

REVIEWED BY
Maria Cristina Crava,
University of Valencia, Spain
Dan-Dan Zhang,
Lund University, Sweden

*CORRESPONDENCE
Ewald Große-Wilde
✉ grosse-wilde@fld.czu.cz

SPECIALTY SECTION
This article was submitted to Chemical Ecology, a section of the journal Frontiers in Ecology and Evolution

RECEIVED 10 October 2022
ACCEPTED 09 December 2022
PUBLISHED 12 January 2023

CITATION
Johny J, Diallo S, Lukšan O, Shewale M, Kalinová B, Hanus R and Große-Wilde E (2023) Conserved orthology in termite chemosensory gene families. *Front. Ecol. Evol.* 10:1065947. doi: 10.3389/fevo.2022.1065947

Conserved orthology in termite chemosensory gene families

Jibin Johny[1], Souleymane Diallo[1], Ondřej Lukšan[2], Mayuri Shewale[1], Blanka Kalinová[1], Robert Hanus[2,3] and Ewald Große-Wilde[1]*

[1]Extemit-K, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Suchdol, Czechia, [2]Chemistry of Social Insects Research Group, Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague, Czechia, [3]EVA 4.0, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Suchdol, Czechia

Termites are eusocial insects known to use a variety of pheromones in tasks necessary for maintenance of their societies. As such, olfaction and pheromone communication in termites has been an object of intense study; trail-following pheromones (TFPs) and sex-pairing pheromones (SPPs), for example, have been identified in many termite species. In contrast, the molecular basis of olfactory detection is understudied in the group. Here, we present chemosensory genes of three species of termites belonging to three distinct lineages, *Neotermes cubanus* (Kalotermitidae), *Prorhinotermes simplex* (Rhinotermitidae), and *Inquilinitermes inquilinus* (Termitidae). Using antennal transcriptome screening of termite workers, we identified the chemosensory genes, which allowed us to perform phylogenetic analysis. We found a comparatively large repertoires of odorant receptors (ORs), gustatory receptors (GRs), ionotropic receptors (IRs), odorant binding proteins (OBPs), chemosensory proteins (CSPs), and sensory neuron membrane proteins (SNMPs). The evolutionary analysis of termite chemosensory genes revealed Isoptera-specific expansions with a 1:1 orthologous pattern, indicating the existence of conserved olfactory functions. Our findings on basal eusocial insects will further enhance our understanding of the molecular underpinnings of eusociality and the evolution of olfactory communication in termites.

KEYWORDS

evolution, olfaction, Blattodea, transcriptome, Isoptera

Introduction

Termites (Blattodea: Isoptera) are the oldest group of eusocial insects; they evolved 140 Ma from wood-dwelling cockroaches within the Blattoidea lineage (Inward et al., 2007; Bucek et al., 2019). Termites display the most pronounced division of labor and caste polyphenism among social insects, manifested by the presence of multiple well-defined caste phenotypes, such as primary and secondary reproductives (kings, queens) and up to several types of workers and soldiers (Roisin and Korb, 2011).

Like other social insects, termites extensively use chemical communication to coordinate the tasks in their colonies, mediate the division of labor, and orient in the environment. Termite chemical communication includes a variety of releaser pheromones directly influencing the behavior of the receivers. These pheromones are used, among others, for marking of foraging trails, mate search, alarm signaling, nestmate recognition, and marking of food sources (Bagnères and Hanus, 2015; Mitaka and Akino, 2021). In addition, termites use primer pheromones having impact on the physiology, reproduction and development of the receivers (Matsuura et al., 2010; Mitaka et al., 2017; Dolejšová et al., 2022).

Trail-following pheromones (TFPs) and sex-pairing pheromones (SPPs) belong among the best studied termite chemical signals. The first TFP has been identified more than five decades ago in the Eastern subterranean termite *Reticulitermes virginicus* (Matsumura et al., 1968). Since then, TFPs have been characterized in 68 species, and SPPs in 17 species of termites, along with a range of other pheromones (primer pheromones, alarm pheromones, etc.; Bordereau and Pasteels, 2010; Bagnères and Hanus, 2015; Mitaka and Akino, 2021). TFPs and SPPs are noteworthy for their conservation and parsimony with respect to chemical diversity. Only 8 different structures occur as TFP across the studied species, some of which are also used as SPPs, or are closely related to the SPPs. At the same time, similarities in chemistry and glandular origin of TFPs and SPPs pinpoint the shared evolutionary origin of trail-following communication and sex-pairing communication (Bordereau and Pasteels, 2010). Experimentally confirmed or expected chemical identities of TFPs and SSPs in the five species studied here – are provided in Table 1.

Like in other insects, antennae are key chemosensory organs in termites (Saran et al., 2007; Du et al., 2019). The antennal flagellum of moniliform antennae is covered by antennal sensilla of various morphological types (Castillo et al., 2021). Irrespective of their different morphologies, all olfactory sensilla house dendrites of olfactory sensory neurons (OSNs), expressing chemosensory receptor proteins that detect the odorants (Hansson and Stensmyr, 2011).

TABLE 1 List of trail-following pheromones (TFPs) and sex-pairing pheromones (SPPs) reported from selected termite species studied here at the level of chemosensory genes.

Termite species	Major pheromone component	Pheromone type
Neotermes cubanus	(3*Z*)-dodecenol*	TFP
Prorhinotermes simplex	(*Z*,*Z*,*E*)-3,6,8-dodecatrien-1-ol	SPP
	Neocembrene and (*Z*,*Z*,*E*)-3,6,8-dodecatrien-1-ol	TFP
Zootermopsis nevadensis	(5*E*)-2,6,10-trimethyl-5,9 undecadienal	SPP - female
	4,6-dimethyldodecanal	SPP and TFP
Reticulitermes speratus	(*Z*,*Z*,*E*)-3,6,8-dodecatrien-1-ol*	SPP
	(*Z*,*Z*,*E*)-3,6,8-dodecatrien-1-ol	TFP

The table is based on the data from Bordereau and Pasteels (2010). *Expected component – based on the pheromone composition reported from congeneric species (Bordereau and Pasteels, 2010).

Odorant Receptors (ORs) and Ionotropic Receptors (IRs) are the two arguably most important protein families involved in the detection of volatile cues in insects (Clyne et al., 1999; Gao and Chess, 1999; Vosshall et al., 1999; Benton et al., 2009; Gomez-Diaz et al., 2018). The ORs and IRs have different structures and are expressed in OSNs associated with distinct olfactory sensillum types (Benton et al., 2009; Scalzotto et al., 2022). Insect ORs are derived from insect Gustatory Receptors (GRs); ORs and GRs form a superfamily with a common phylogenetic origin (Robertson et al., 2003; Thoma et al., 2019). ORs are seven-transmembrane receptors that exhibit unusual topology compared to classical GPCRs, with an intracellular N-terminus. OR proteins function as heteromultimeric channels formed by a ligand-specificity defining OR protein, and the ubiquitous OR-Coreceptor (ORCo; Vosshall et al., 1999; Hansson and Stensmyr, 2011; Butterwick et al., 2018). Silencing ORCo in two termite species (*Reticulitermes chinensis* and *Odontotermes formosanus*) impaired their ability to perceive TFPs and perform oriented locomotory behavior, indicating a role of ORs in termite pheromone detection (Gao et al., 2020). ORs exhibit extreme sequence variability, with rapid birth-and-death evolution leading to gene diversifications (McBride and Arguello, 2007; Sánchez-Gracia et al., 2009; Ramdya and Benton, 2010).

ORs can detect a wide range of environmental odors. The remarkable diversity of potential chemical cues, whose importance changes in conjunction with lifestyle changes across insects, presumably drives the rapid birth-and-death evolution seen in ORs (McBride and Arguello, 2007; Sánchez-Gracia et al., 2009; Ramdya and Benton, 2010). Accordingly, OR repertoire sizes differ significantly among insect species, with for example 62 ORs reported in *Drosophila melanogaster* (Clyne et al., 1999; Gao and Chess, 1999; Vosshall et al., 1999), 131 in mosquitoes (Bohbot et al., 2007), 77 in bark beetles and as many as 375 ORs in the ant *Acromyrmex echinatior* (Zhou et al., 2015). The evolution of this gene family is clearly driven by the diversity of chemical recognition needs, and the evolution of eusocial organization in social Hymenoptera (bees, ants, wasps) shows correlation with the diversification of OR repertoire (Yan et al., 2020).

Gustatory receptors (GRs) are seven-transmembrane domain proteins that are first identified in *D. melanogaster*, involved in the detection of tastants: sugar, bitter and CO_2 (Clyne et al., 2000; Scott et al., 2001). Together with ORs they form a superfamily of proteins. They share the same inverted topology when compared to classical G-protein coupled receptors (Scott et al., 2001; Zhang et al., 2011). Molecular evolutionary analysis has proposed GRs as ancestral sequences to ORs (Robertson et al., 2003). A number of GRs have been functionally characterized in *D. melanogaster*, DmelGR32a and DmelGr68a are involved in pheromone detection

(Montell, 2009), and DmelGr21a and DmelGr63a, which function together as CO_2 receptor (Jones et al., 2007). Various other fly receptors are involved in the detection of sugars and bitter tastants (Dahanukar et al., 2001, 2007; Vosshall and Stocker, 2007; Montell, 2009; Isono and Morita, 2010; Freeman et al., 2014; Delventhal and Carlson, 2016).

IRs are transmembrane proteins distantly related to a variant of ionotropic glutamate receptor receptors, iGluRs, expressed with co-receptors IR8a or IR25a (Benton et al., 2009; Croset et al., 2010). In general, iGluRs are classified into subfamilies defined by their main agonist: α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA), kainate, and N-methyl-D-aspartate (NMDA; Croset et al., 2010). AMPA and kainate receptors were further verified by the presence of threonine (T) in the first half of the S2 domain; iGluRs that lack this T residue were classified as NR1 or delta variants (Benton et al., 2009). IRs lack the conserved aspartate (D) or glutamate (E) that interacts with the α-amino group of the glutamate ligand, that non-IR iGluRs possess (Benton et al., 2009). Like other iGluRs, IRs function as heterotetrameric channels; they participate in various sensory modalities, including olfaction. Antennal IRs are considerably more conserved than ORs, with a similarly conserved ligand profile (Benton et al., 2009; Abuin et al., 2019).

Additionally, other non-receptor proteins like odorant binding proteins (OBPs), chemosensory proteins (CSPs), and sensory neuron membrane proteins (SNMPs; Hansson and Stensmyr, 2011; Leal, 2011) are involved in olfaction (Zhou et al., 2006). OBPs and CSPs are both small, globular proteins that are present in high concentrations within the sensillum lymph (Venthur and Zhou, 2018; Pelosi et al., 2018b) and both bind odorant molecules (Pelosi et al., 2005). Their exact function is so far unclear, but there is evidence that they might be involved in the transport of hydrophobic odorants through the lymph space to the dendrite membranes, or that they might protect odorants from enzymatic degradation (Pelosi et al., 2018a). SNMPs belong to the CD36 family of transmembrane proteins; their exact function in olfaction is unclear, but functional knock-down in *D. melanogaster* indicated that they are involved in pheromone detection (Benton et al., 2007; Pregitzer et al., 2014).

Chemosensory genes have been identified in a wide range of insects using genomic and transcriptomic approaches (Clyne et al., 1999; Gao and Chess, 1999; Vosshall et al., 1999) which to a large extent enabled us to understand their evolutionary and behavioral adaptations in different contexts of biology including eusociality (Robertson and Wanner, 2006; Zhou et al., 2012, 2015; Engsontia et al., 2014; Terrapon et al., 2014; De Fouchier et al., 2017; Pask et al., 2017; Auer et al., 2020; Obiero et al., 2021; Keesey et al., 2022). Termites, despite being eusocial insects with well-studied chemical ecology and pheromone biology, have not been examined in detail in this regard until recently, with chemosensory genes identified only in three species: *Zootermopsis nevadensis* (Archotermopsidae) with 85 ORs in the genome (Terrapon et al., 2014), *Cryptotermes secundus* (Kalotermitidae) with 42 ORs in the genome (Harrison et al., 2018) and *Reticulitermes speratus* (Rhinotermitidae) with 22 ORs in the whole-body transcriptome (Mitaka et al., 2016). The relatively modest repertoire of termite ORs contrasts with the situation in eusocial Hymenoptera and the assumptions made on the impact of social evolution on OR expansion. At the same time, it highlights the independent origin of eusociality in Hymenoptera and Isoptera, and may potentially reflect the chemical parsimony observed in termite semiochemicals. By contrast, the repertoire of termite IRs is substantially expanded, with 141 and 135 IRs reported, respectively, for *Z. nevadensis* (Terrapon et al., 2014) and *C. secundus* (Harrison et al., 2018), which may intuitively be interpreted as a correlate of the social lifestyle (Harrison et al., 2018).

Here, we present the results of antennal transcriptome analysis of termite workers for the identification of the main olfactory sensory genes of three termite species, *Neotermes cubanus* (Kalotermitidae), *Prorhinotermes simplex* (Rhinotermitidae), and *Inquilinitermes inquilinus* (Termitidae). The three species were selected to cover the phylogenetic diversity of Isoptera, from the relatively basal Kalotermitidae to the modern lineage of Termitidae, with Rhinotermitidae being situated on the mid-way between the two. At the same time, the life histories of the three species represent different levels of social complexities encountered in termites, from the socially primitive *N. cubanus* devoid of true worker caste through *P. simplex* situated at the boundary of the emergence of true foraging workers, to the socially advanced higher termite *I. inquilinus* having true worker caste (Rupf and Roisin, 2008; Roisin and Korb, 2011). Our findings should serve as a basis in elucidating the evolution of pheromone detection in termites.

Materials and methods

Insect origin and antennal tissue dissection

We used the following termite species for RNA sequencing and *de novo* assembly of antennal transcriptome: *N. cubanus*, *P. simplex*, and *I. inquilinus*. Multiple colonies of *N. cubanus* (Snyder) and *P. simplex* (Hagen), originating from field collections in Cuba, are kept in laboratory at the Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences. Colonies live in glass vivaria at 27°C and 80% relative humidity in clusters of spruce wood slices. Mature colony of *I. inquilinus* (Emerson) was collected by the authors during the field mission to French Guiana along the Road to Petit Saut (N05 03.975 W053 02.764) in 2019 with the consent of Office National des Forêts (Cayenne).

Ninety workers from one colony per species were cold-anesthetized, quickly washed in cold ethanol and decapitated under a stereomicroscope. Heads were transferred into RNase free collection tubes and kept at 4°C overnight in 1 ml of RNA*later* (Thermo Fisher Scientific). Antennae were dissected the next day,

collected in a droplet of 96% ethanol, snap frozen and stored at −80°C until RNA extraction.

RNA extraction and sequencing

Total RNA from pools of 180 worker antennae from each species was extracted using acid guanidinium thiocyanate-phenol-chloroform extraction. Deep frozen samples were transferred on liquid nitrogen, grinded using PP pestle directly in the collection tube and homogenized at room temperature after addition of TRI reagent solution (Thermo Fisher Scientific). Extraction steps included vortexing and centrifugation (15,000 *g*, 15 min at 4°C), RNA precipitation using isopropanol (1:1, followed by centrifugation 15,000 *g*, 15 min at 4°C), washing in 75% ethanol (centrifugation 5,000 *g*, 5 min at 4°C), drying at room temperature in laminar flow box and resuspension in 10 mM Tris–HCl, pH 8.0 with 0.1 mM EDTA. The quality and quantity of isolated RNA was inspected on Nanodrop ND-1000 UV/VIS spectrophotometer and Qubit 4 fluorometer using the RNA HS Assay Kit (all Thermo Fisher Scientific), the integrity was evaluated on 1% agarose gel after staining with ethidium bromide.

Library preparations of all three antennal poly(A)-selected strand-specific cDNA libraries and high throughput sequencing analysis on Illumina HiSeq with 30 millions of 2 × 150 paired end reads was conducted at Eurofins Genomics (Ebersberg, Germany).

Transcriptome assembly and gene annotation

Raw sequencing reads were inspected for erroneous k-mers and corrected with rCorrector (Song and Florea, 2015), residual sequencing adapters and low-quality bases were trimmed using Trimmomatic v0.32 (Bolger et al., 2014). Sequence contaminants, such as ribosomal RNA, were filtered out based on mapping to the reference from SILVA database release 132 (Quast et al., 2013) with bowtie2 v2.3.4.1 algorithm (Langmead and Salzberg, 2012) or by depletion of overrepresented sequences using the RemoveFastqcOverrepSequenceReads.py script.[1] *De novo* assembly of antennal transcriptomes was performed with Trinity v2.1.1 in default settings for strand-specific reads (Grabherr et al., 2011), candidate coding regions were identified upon prediction of open reading frames with Transdecoder v5.5.0.[2] The raw data used for transcriptome assembly are deposited in the NCBI SRA repository with BioSample accession numbers: SAMN31093778 (*Inquilinitermes inquilinus*), SAMN31093779 (*Neotermes cubanus*) and SAMN31093780 (*Prorhinotermes simplex*).

To assess the completeness of the transcriptomes, BUSCO 5.3.2 (Simão et al., 2015) was used to test for the presence of the Insecta odb10 reference genes within the transcriptome assemblies. For manual annotation, we created databases based on the longest assembled isoform of each transcript. BLASTx searches (Camacho et al., 2009) were performed on these local databases using reference datasets of each multigene family: ORs, IRs, OBPs, CSPs and SNMPs as queries with an e-value cut-off of 0.001. For ORs, reference datasets included amino acid sequences of *Z. nevadensis* (nr), *C. secundus* (nr), and *D. melanogaster* (Refseq NCBI) as well as sequences of *Ampulex compressa*, *Cerceris arenaria*, *Psenulus fuscipennis*, *Apis mellifera*, *Bombus terrestris*, *Habropoda laboriosa*, *Dufourea noveangliae*, *Lasioglossum albipes*, *Nasonia vitripennis*, *Harpegnathus saltator*, and *Solenopsis invicta* previously reported in Obiero et al. (2021). OR candidate protein sequences were further subject to analysis for presence of the correct transmembrane domains using TMHMM 2.0 (Krogh et al., 2001). In the final step, prior to multiple sequence alignment and phylogenetic analyses, all sequences with insufficient similarities comparing to the reference dataset were manually filtered out based on an all-against-all BLAST analysis and subsequent clustering in CLANS (Frickey and Lupas, 2004). GR candidates were predicted from the termite transcriptomes using a reference dataset containing GR amino acid sequences from *Z. nevadensis* (Terrapon et al., 2014), *C. secundus* (Harrison et al., 2018), *R. speratus* (Mitaka et al., 2016), *D. melanogaster* (nr), and *Tribolium castaneum* (NCBI RefSeq).

For IRs, we used sequences reported from *Z. nevadensis* (Terrapon et al., 2014) and *C. secundus* (Harrison et al., 2018) and the iGluR amino acid sequences from Croset et al. (2010). For OBPs and CSPs the reference data set included amino acid sequences from *Z. nevadensis* (nr), *C. secundus* (nr), *D. melanogaster*, and *Locusta migratoria*, as well as the dataset used in Vogt et al. (2015) and Guo et al. (2018). The SNMP dataset was created using sequences from the termites *Z. nevadensis* (nr), *C. secundus* (nr), the fruit fly *D. melanogaster* (nr), the beetle *Aethina tumida* (nr), the moth *Manduca sexta* and SNMPs from three Coleopteran species *T. castaneum* (nr), *R. palmarum* (Gonzalez et al., 2021), *Sitophilus oryzae* (nr). Finally, the predicted amino acid sequences of each multigene protein family were retrieved manually from the transcriptome assemblies based on the blastx search results. Additionally, in the odorant binding proteins, signal peptides were predicted using SignalP v6.0 (Teufel et al., 2022).

Phylogenetic analysis of the candidate chemosensory proteins

The phylogenetic reconstruction of each protein family was performed using the Maximum Likelihood method (Felsenstein, 1981). To compare and predict phylogenetic relationships, we retrieved available relevant chemoreceptor protein sequences from the GenBank nr database. For the termite odorant receptor phylogeny, we used OR sequences from the termite species

1 https://github.com/harvardinformatics/TranscriptomeAssemblyTools

2 https://github.com/TransDecoder

Z. nevadensis, *C. secundus*, *N. cubanus*, *P. simplex*, *R. speratus*, and *I. inquilinus*, as well as the termite relative, the cockroach *Blattella germanica*. We further included *Bombyx mori* and *Manduca sexta* (Lepidoptera), *Drosophila melanogaster* (Diptera), *Ips typographus* and *Tribolium castaneum* (Coleoptera), *Forficula auricularia* (Dermaptera), *Athalia rosae*, and *Apis mellifera* (Hymenoptera). As an outgroup, the crustacean *Daphnia pulex* Gr 42, 43, and 44 (Saina et al., 2015) sequences were used. The larger number of datasets was required to reach a predicted phylogeny with sufficient support, likely due to the high sequence diversity of ORs. Multiple sequence alignment was performed using MAFFT v.7 (Katoh et al., 2017) under the E-INS-i iterative refinement method, followed by trimming using trimAl v1.4 (Capella-Gutiérrez et al., 2009) with the "automated1' option. The same alignment and trimming methods were followed for the other chemosensory gene families. The best-fit amino acid substitution model, JTT + G + F was determined for ORs using ProtTest v.3.4.2 (Darriba et al., 2017) under AIC criteria. Phylogenetic reconstructions of all chemosensory genes were performed by means of the maximum likelihood method using RAxML-NG 1.1.0 (Kozlov et al., 2019) and 1,000 bootstrap replications.

The phylogeny of gustatory receptors was reconstructed using JTT + F + G4 as the best-fit amino acid substitution model and rooted with CO_2 and sugar receptors as their basal location was reported earlier in analyses with GRLs of other animals (Robertson, 2015; Robertson et al., 2018). Gustatory receptors reported from the termite species *N. cubanus*, *P. simplex*, *I. inquilinus*, *Z. nevadensis*, *C. secundus*, and *R. speratus* were used in the analysis. Additionally, GRs from the non-blattodean species *D. melanogaster* and *T. castaneum* were included in the analysis allow classification of different subclades including receptors for CO_2 and receptors for bitter and sweet tastes.

The ionotropic receptor phylogeny was reconstructed using LG + F + R as the best-fit amino acid substitution model and was rooted with non-NMDA iGluRs as an outgroup, using 1,000 replicates to calculate bootstrap support. The amino acid sequences from the following species were added to study the phylogenetic relationship: the termites *Z. nevadensis*, *C. secundus*, *N. cubanus*, *P. simplex*, *I. inquilinus*, the cockroach *B. germanica*, the fruit fly *D. melanogaster*, and the beetles *Dendroctonus ponderosae* and *Rhynchophorus palmarum*. Non-blattodean species were added to allow for better determination of the correct iGluR-subclades of novel candidates.

The SNMP phylogeny was reconstructed using LG + R as the best-fit amino acid substitution model under Bayesian information criterion with 1,000 bootstrap replications. Species compared in the phylogeny were the termites *Z. nevadensis*, *C. secundus*, *N. cubanus*, *P. simplex*, and *I. inquilinus*, the cockroach *B. germanica*, the fruit fly *D. melanogaster*, the beetles *T. castaneum*, *Sitophilus oryzae* and *R. palmarum*, the moths *B. mori* and *M. sexta*, and the ant *Harpegnathos saltator.* Coleopteran SNMPs are included in the phylogeny as additional SNMP groups are reported in this insect order (Dippel et al., 2016; Zhao et al., 2020). We used *D. melanogaster* Croquemort (crq) protein, a member of the CD36 family but not an SNMP, as an outgroup.

The maximum likelihood phylogeny of termite OBPs was reconstructed using LG + R as the best-fit amino acid substitution model under AIC with 1,000 bootstrap replications. Bristletail *Lepismachilis y-signata* OBPs were used as outgroup. The species included in the analysis were the termites *Z. nevadensis*, *C. secundus*, *N. cubanus*, *P. simplex*, and *I. inquilinus*, the beetles *T. castaneum* and *R. palmarum*, the moth *M. sexta* and the fruit fly *D. melanogaster.* The Maximum likelihood phylogeny of termite CSPs was constructed using LG + R as amino acid substitution model and rooted with *D. pulex* CSP sequences as outgroup. The other species included in the analysis were the termite *R. speratus*, the beetles *T. castaneum* and *R. palmarum*, the moth *B. mori*, the honey bee *Apis mellifera*, the fruit fly *D. melanogaster*, the chironomid *Clunio marinus*, and the ant *Camponotus japonicus.* Inclusion of CSPs from the listed species allowed a better comparison of termite CSPs across insect orders.

Results

De novo antennal transcriptome sequencing and assembly

We generated antennal transcriptome data for *N. cubanus*, *P. simplex*, and *I. inquilinus* from Illumina paired-end sequencing. This yielded 48.0 million read pairs from *N. cubanus* libraries, resulting in 247,031 transcripts, with a total of 53,949 predicted ORFs based on Trinity *de novo* assembly. The same approach generated 46.3 million read pairs, yielding 180,250 transcripts with 58,126 predicted ORFs in *P. simplex*, and 30.6 million read pairs assembled into 203,568 transcripts that included 52,980 predicted ORFs in *I. inquilinus*. Next, we performed BUSCO 5.3.2 analysis as a measure for completeness of the transcriptomes, using the insecta10 dataset as a reference. This analysis showed 97.4, 97.3 and 97.2% completeness for *N. cubanus*, *P. simplex*, and *I. inquilinus*, respectively. An overview of the sequencing and assembly statistics is provided in Table 2.

Termite odorant receptors

We manually annotated sequences of the transcriptome assemblies coding for members of the major insect chemosensory families, starting with olfactory receptors (ORs). We recovered 30, 50 and 28 ORs from the antennal transcriptomes of *N. cubanus*, *P. simplex*, and *I. inquilinus*, respectively. Of these, 24, 48 and 27 predicted proteins, respectively, presented an OR-typical transmembrane profile in TMHMM analysis (Krogh et al., 2001), and a length of >350 aa, which we considered to be full-length ORs. Next, we reconstructed a maximum likelihood phylogeny using the predicted amino acid sequences of our candidate ORs from the three studied species, as well as other

TABLE 2 *De novo* transcriptome assembly statistics of the three species of termites, *N. cubanus*, *P. simplex*, and *I. inquilinus*.

	Neotermes cubanus	*Prorhinotermes simplex*	*Inquilinitermes inquilinus*
Total number of raw reads	48,033,206	46,252,432	30,635,256
Number of transcripts	247,031	180,250	203,568
Full length ORFs	53,949	58,126	52,980
N50 length	2,273	3,482	2,356
GC content	39.57	39.55	40.19
Complete BUSCOs (insectaodb_10, %)	97.4	97.3	97.2
Fragmented BUSCOs (insectaodb_10, %)	1.0	1.0	0.6

termite-ORs that were previously reported from *Z. nevadensis* (Terrapon et al., 2014), *C. secundus* (Harrison et al., 2018) and *R. speratus* (Mitaka et al., 2016). We also added OR coding sequences of *B. germanica* (Robertson et al., 2018), as well as ORs from other major insect orders, to help stabilizing the phylogenetic analysis and assist in the examination of our newly identified termite ORs. To add more resolution to the phylogeny, we also added a set of recently reported 'primitive ORs' from the silverfish *Lepisma saccharina* (Thoma et al., 2019). Finally, we included gustatory receptors from *D. pulex* that had previously been shown to be an outgroup for all insect ORs (Pēalva-Arana et al., 2009).

The phylogeny, rooted using the *D. pulex* GR outgroup, revealed the monophyly of OR and GR gene families with high bootstrap support. Between these two major clades was a group of 'GR and OR-like' sequences representing mainly termites and Lepidoptera. Adding a larger number of non-isopteran sequences, including ORs of the basal insect *L. saccharina* helped stabilising the phylogeny of this clade of GR and OR-like sequences, with isopteran sequences sharing more sequence similarity with GRs than with the highly expanded OR families across other insect orders. We found representatives from *Z. nevadensis*, *C. secundus*, and *P. simplex* within this clade, but not from other termite species or *B. germanica*. The ORs of Zygentoma formed an ancestral clade with high bootstrap support, and within this, ORCos appeared as highly derived sequences, fitting reports of the evolutionary origin and ancestral nature of ORCo sequences (Missbach et al., 2014; Brand et al., 2018; Thoma et al., 2019). The termite ORCos formed a subset with primitive ORCo from *L. saccharina* as an ancestral sequence. The three Isoptera-specific expansions in ORs are in accordance with the other insect orders and indicate an evolutionary pattern, i.e., an ancestral set of ORs that share orthologous sequences between most insect orders, and a rapidly evolving set with multiple species-specific expansions, as mainly observed in Coleoptera and Hymenoptera (Andersson et al., 2019). We found 13 isopteran ORs in the ancestral clade, sharing orthologs to different insect orders and the remaining ORs formed two independent Isoptera-specific expansions of 50 ORs and 37 ORs. Within these two expansions, the most recently evolved one (37 ORs) shares sequence similarity with hymenopteran ORs whereas the other one (50 ORs) was similar to the ancestral isopteran clade (Figure 1).

Termite gustatory receptors

Our manual annotation revealed 20, 25 and 26 GRs, respectively, from the antennal transcriptomes of *N. cubanus*, *P. simplex*, and *I. inquilinus*. Several of these sequences already were identified in the search for ORs, and have been labelled as GRs in Figure 1. Among these candidate genes, 8, 7 and 6 receptors, respectively, from *N. cubanus*, *P. simplex*, and *I. inquilinus* belong to the clade containing the *D. melanogaster* CO_2 receptor clade. Similarly, 6, 3, and 4 candidates, respectively, from *N. cubanus*, *P. simplex*, and *I. inquilinus* belong to the sugar receptor clade, and 4, 15 and 16 candidates, respectively, to the bitter taste receptor clade (Figure 2). Isoptera- specific expansions were observed in all three subclades. However, we found no clear 1:1 orthologous for the *D. melanogaster* pheromone sensitive or CO_2 GRs in any of the Isopteran GRs compared. Putative orthologs for the *D. melanogaster* fructose receptor Gr43a were present in *Z. nevadensis*, *C. secundus*, and *T. castaneum*.

Isoptera-specific expansions in termite antennal ionotropic glutamate receptors

Next, we analysed putative IRs coding sequences. BLASTx searches were performed using well-annotated IR and iGluR sequences from different insect orders, Coleoptera, Lepidoptera, Hymenoptera and Diptera (Croset et al., 2010). Using this approach, we recovered, 98, 95 and 77 transcripts from the antennal transcriptomes of *N. cubanus*, *P. simplex*, and *I. inquilinus*, respectively. Based on length of the predicted protein, as well as presence of all IR-typical domains, we considered 33, 53 and 29 transcripts from *N. cubanus*, *P. simplex*, and *I. inquilinus*, respectively, as complete. In multiple sequence alignment, we confirmed iGluRs family members by the presence of a characteristic conserved arginine (R) residue in the S1 domain involved in binding the glutamate α-carboxyl group (Benton et al., 2009; Croset et al., 2010). We further classified these receptors into the three distinct iGluR subfamilies (AMPA, NMDA, kainate) based on homology.

Finally, IR subfamily members were identified based on the absence of conserved aspartate (D) or glutamate (E) in the second half of the S2 domain that interacts with the α-amino

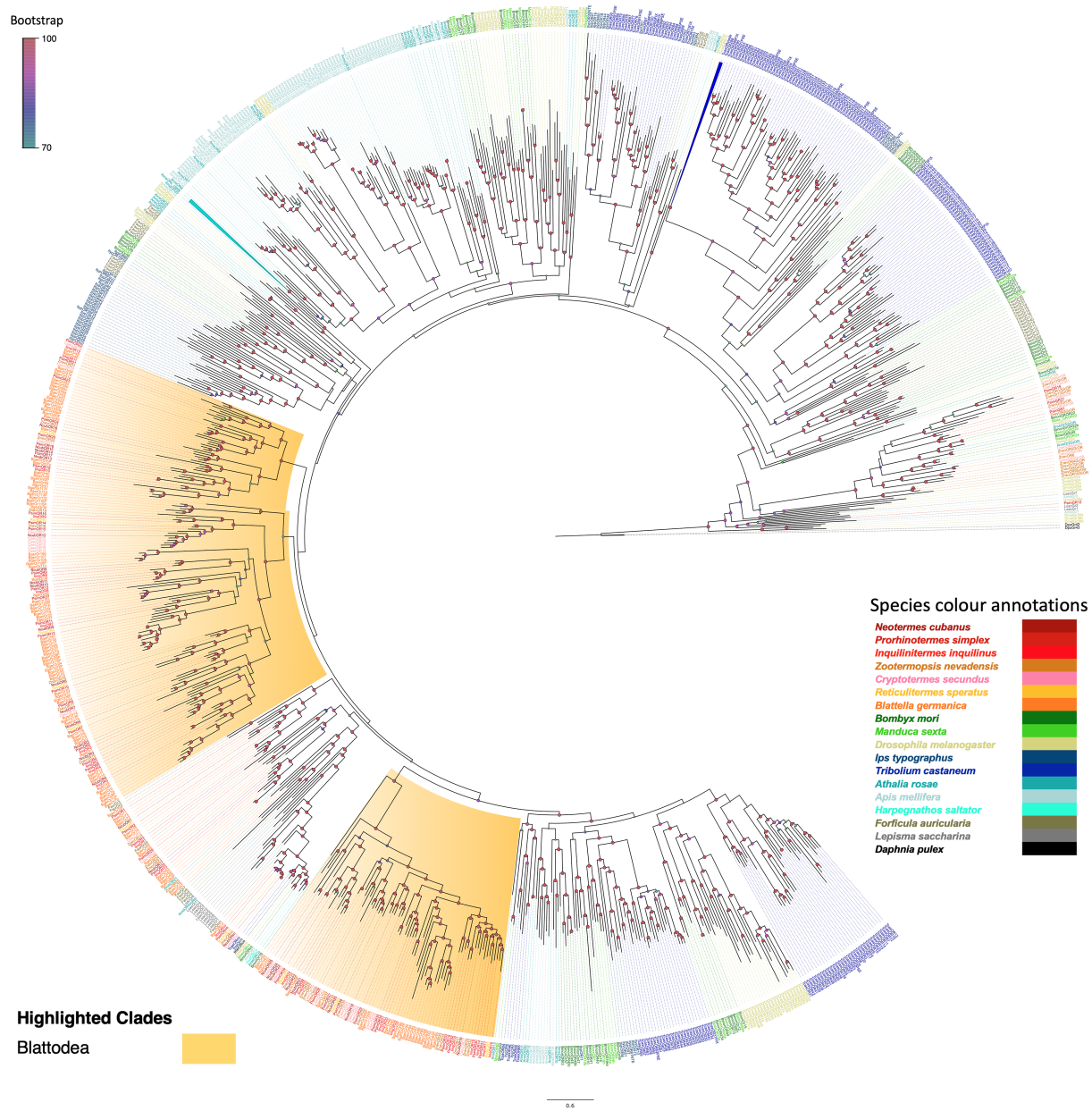

FIGURE 1
Maximum likelihood phylogeny of termite ORs. The tree was reconstructed using predicted OR sequences from termites and other insect orders and rooted using GRs as an outgroup. ORs from other species were named as follows: Blattodea: *R. speratus* (Rspe), *C. secundus* (Csec), *N. cubanus* (Ncub), *P. simplex* (Psim), *I. inquilinus* (Iinq), *Z. nevadensis* (Znev), and *B. germanica* (Bger); Lepidoptera: *B. mori* (Bmori) and *M. sexta* (Msex); Diptera: *D. melanogaster* (Dmel); Coleoptera: *I. typographus* (Ityp) and *T. castaneum* (Tcas), Dermaptera: *F. auricularia* (Faur) and Hymenoptera: *A. rosae* (Aros) and *A. mellifera* (Amel), *H. saltator* (Hsal); Outgroup: *D. pulex* (Dpu) and primitive ORs: *L. saccharina* (Lsac). Sequence names are colored according to species and color codes are provided in the in-figure legend. Node color indicates the bootstrap support value based on 1,000 replicates. The scale bar indicates the estimated amino acid substitutions per site.

group of the glutamate ligand. Partial sequences that were too short to include these protein domains were excluded from the analysis, but classified based on homology alone. For the phylogenetic analysis, we used our newly identified iGluR sequences, as well as other termite iGluR and IR sequences that were previously reported from *Z. nevadensis* (Terrapon et al., 2014) and *C. secundus* (Harrison et al., 2018). In some cases, partial sequences we identified were too short to reasonably include in phylogenetic and where therefore exluded, most importantly the IR8a candidate of *I. inquilinus*. We also added iGluR coding sequences from other major insect orders to stabilize the analysis and assist in the annotation of newly identified termite IRs and iGluRs. Finally, we used non-NMDA iGluRs from the *D. pulex* as an outgroup as these receptors are

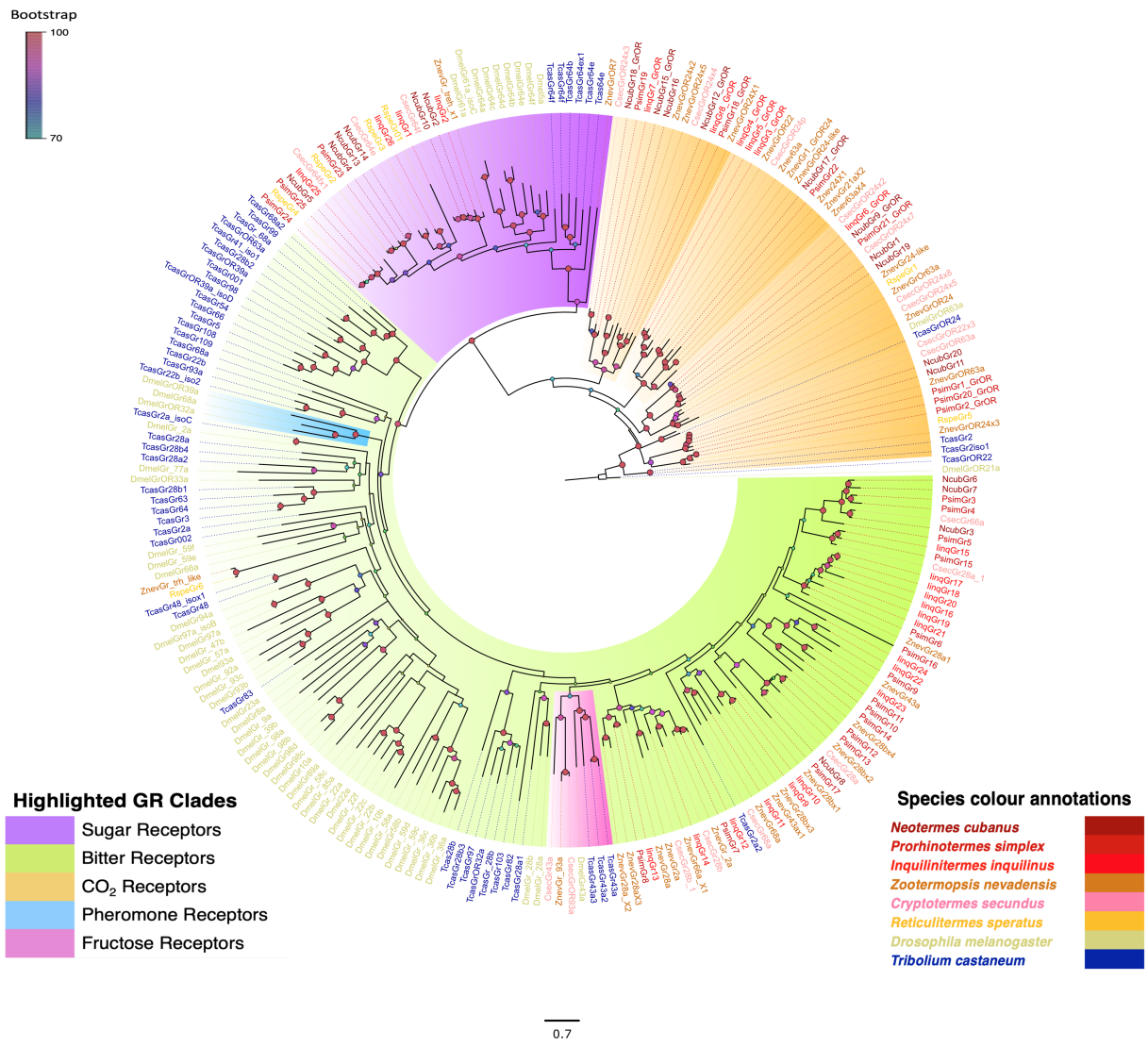

FIGURE 2
Maximum likelihood phylogeny of termite GRs highlighting known taste and CO_2 receptors. The highlighting details are provided in-figure legend. The tree was constructed using JTT+F+G4 as best-fit amino acid substitution model and rooted with CO_2 and sugar receptors as their basal location reported earlier in analyses with GRLs of other animals (Robertson, 2015; Robertson et al., 2018). Species included in the phylogeny are *N. cubanus* (Ncub), *P. simplex* (Psim), *I. inquilinus* (linq), *Z. nevadensis* (Znev) *R. speratus* (Rspe), *C. secundus* (Csec), *T. castaneum* (Tcas) and *D. melanogaster* (Dmel). Sequence names are colored according to species and color codes are provided in the in-figure legend. Node color indicates the bootstrap support value based on 1,000 replicates. The scale bar indicates the estimated amino acid substitutions per site.

considered to be ancestral to both NMDA iGluRs and IRs (Croset et al., 2010).

After rooting, the dendrogram revealed clear monophyletic clades for each major iGluR subfamily with maximum bootstrap support (Figure 3). The non-NMDA iGluR subfamilies appeared basal in the phylogeny, with the IR8a, IR25a, and the NMDA clades highly derived. We found representative sequences from all three clades in all three termite transcriptomes with IinqIR8a being partial sequence excluded from the analysis. The remaining iGluRs formed three separate clades; the first two were grouped as antennal IRs (including an Isoptera-specific subclade), and the third one as divergent IRs based on the classification scheme used in *Drosophila* IRs (Benton et al., 2009). The three species shared a nearly equal number of antennal IRs (*N. cubanus*: 36, *P. simplex*: 37, *I. inquilinus*: 34), i.e., 36–44% of total IRs identified. Within antennal IRs, 13, 19 and 14 transcripts from *N. cubanus*, *P. simplex* and *I. inquilinus*, respectively, formed an Isoptera-specific clade (Figure 3). These numbers were on par or slightly higher than previously reported for other termite species for example, 12 IRs reported from *R. speratus* transcriptome (Mitaka et al., 2016). The clade of divergent IRs showed weak bootstrap support. However, the Isoptera-specific expansion of both antennal and divergent IRs was well supported (Figure 3).

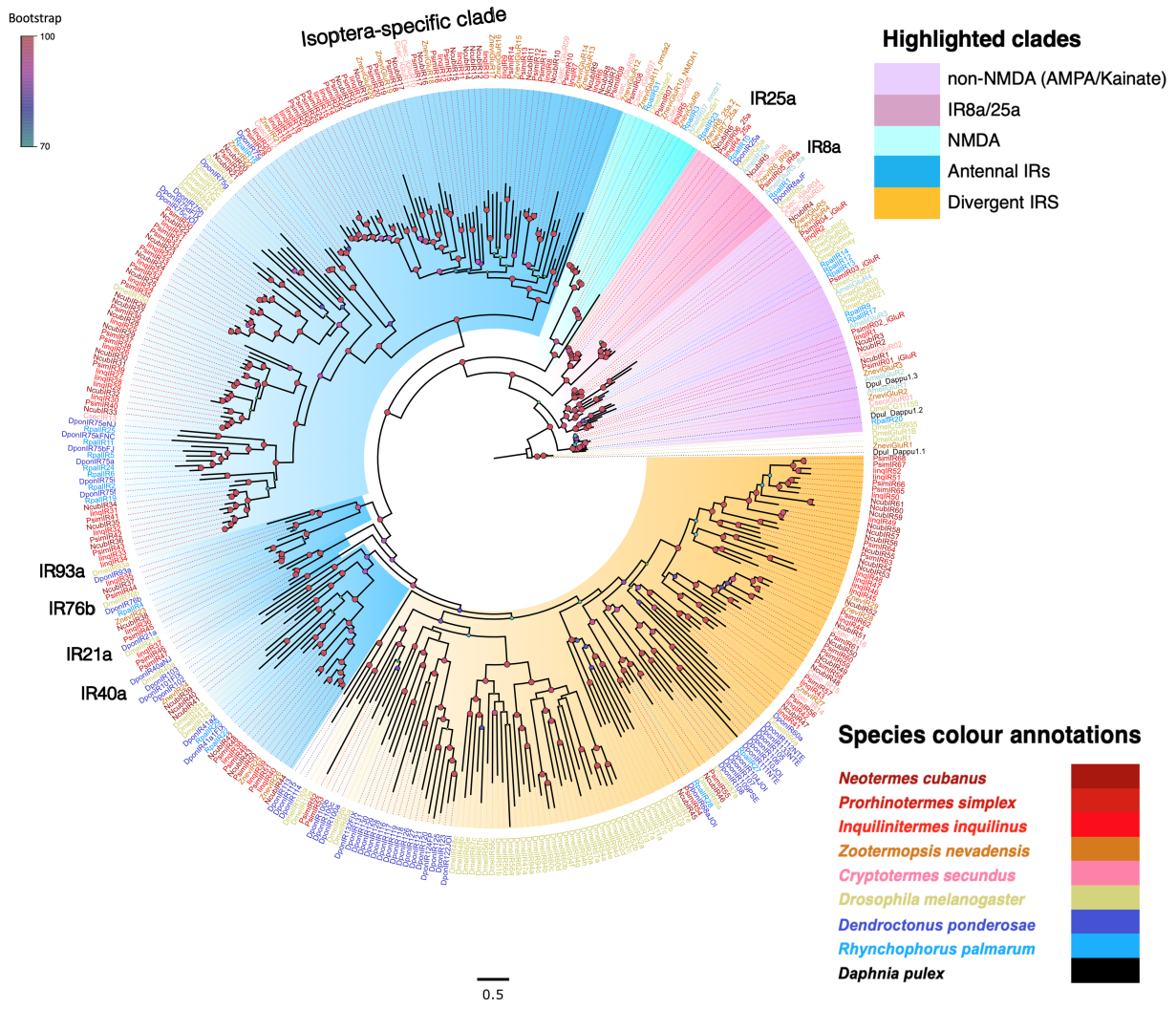

FIGURE 3
Phylogeny of termite IRs showcasing major iGluR subfamilies. The tree was reconstructed using the maximum likelihood method using LG+F+R as the best-fit amino acid substitution model and was rooted with non-NMDA iGluRs as an outgroup. Major iGluR subfamilies were highlighted as non-NMDA iGluRs (Purple), IR8a/25a (pink), NMDA (light blue), antennal IRs (blue), Divergent IRs (orange). Known conserved IR subgroups are marked in the tree. Each species represented were named as *C. secundus* (Csec), *N. cubanus* (Ncub), *P. simplex* (Psim), *I. inquilinus* (Iinq), *Z. nevadensis* (Znev), *D. melanogaster* (Dmel), *D. ponderosae* (Dpon), *R. palmarum* (Rpal), and *D. pulex* (Dpul). Sequence names are colored according to species and color codes are provided in the in-figure legend. The node colors indicate the bootstrap support (1,000 replicates) and the scale represents the estimated amino acid substitutions per site.

Sensory neuron membrane proteins in termite antennal transcriptome

The next family of chemosensory proteins investigated were sensory neuron membrane proteins (SNMPs). BLASTx query using well-annotated sequences of SNMP1 and SNMP2 recovered six transcripts each from *I. inquilinus* and *P. simplex* and five from *N. cubanus* as SNMPs. We added previously reported SNMPs from *Z. nevadensis* (Terrapon et al., 2014), *C. secundus* (Harrison et al., 2018) and SNMPs from other insect orders to our data for phylogenetic analysis using maximum likelihood algorithms. Additionally, we used a non-SNMP CD36 family protein, croquemort (crq) from *D. melanogaster* as outgroup (Figure 4). Based on the phylogeny, we identified 5 out of 6 transcripts each from *P. simplex* and *I. inquilinus*, and 4 out of 5 from *N. cubanus* as SNMP1. We found four Isoptera-specific SNMP1 subclades with high bootstrap support and thus, further classification in subtypes 'a' and 'b' as in other orders was not attempted. We identified one SNMP2 protein each from all the three-termite species analysed, which also formed an Isoptera-specific clade.

Soluble proteins (OBPs and CSPs) involved in termite chemoreception

Starting with well-annotated sequences from other insect species we also screened our transcriptomes for sequences encoding candidate OBPs. Using this approach, we recovered

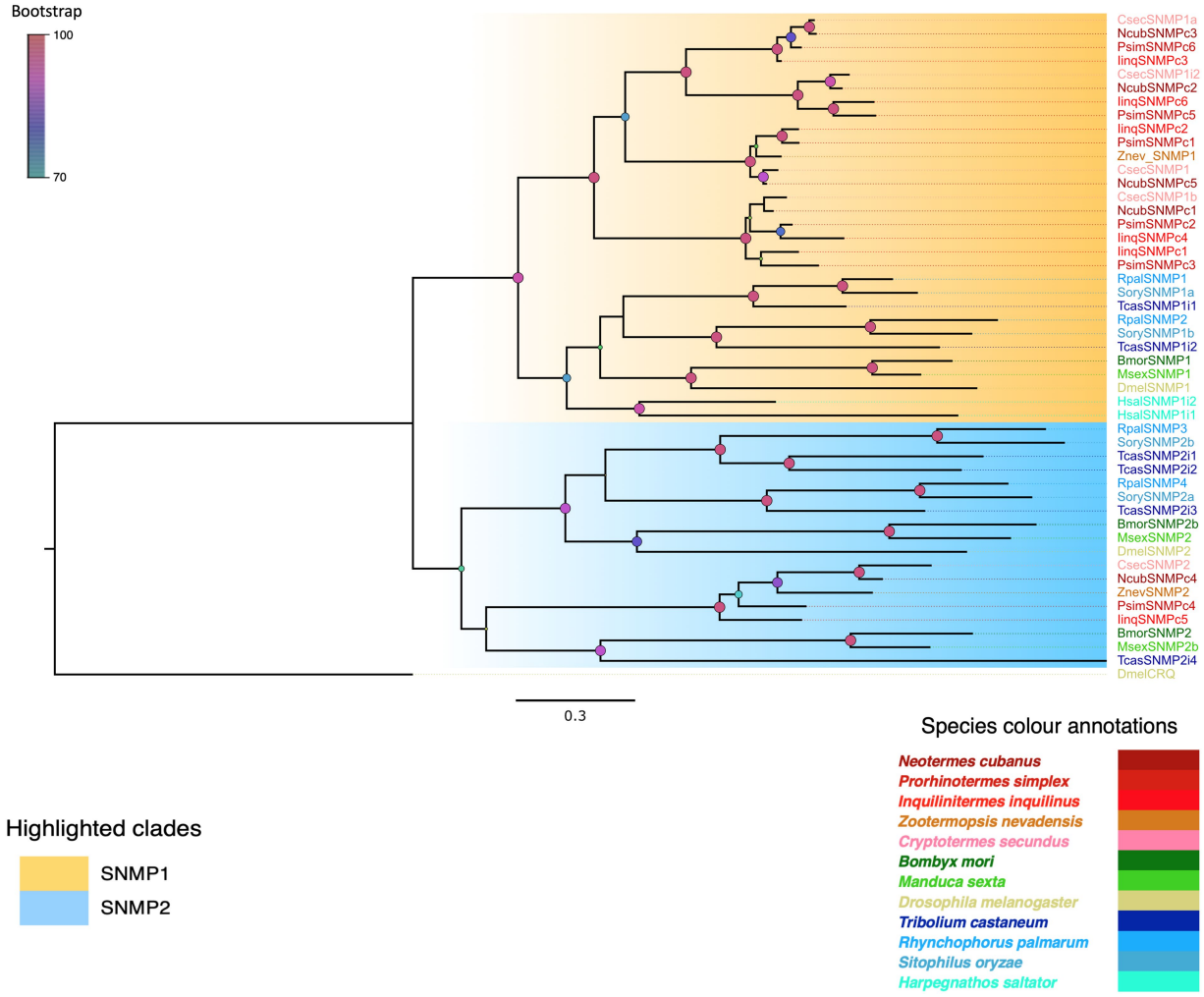

FIGURE 4
Maximum likelihood phylogenetic tree indicating two SNMP types in termites. The tree was constructed using LG+R as the best-fit amino acid substitution model under Bayesian information criterion. The two SNMP types are highlighted as SNMP1 (orange) and SNMP2 (blue). Species abbreviations used are *C. secundus* (Csec), *N. cubanus* (Ncub), *P. simplex* (Psim), *I. inquilinus* (linq), *Z. nevadensis* (Znev), *D. melanogaster* (Dmel), *S. oryzae* (Sory), *T. castaneum* (Tcas), *R. palmarum* (Rpal), *B. mori* (Bmor) *M. sexta* (Msex), and *H. saltator* (Hsal). Sequence names are colored according to species and color codes are provided in the in-figure legend. The node colors indicate the bootstrap support (1,000 replicates). The scale bar represents the estimated amino acid substitutions per site.

29, 34 and 25 candidates from *N. cubanus*, *P. simplex*, and *I. inquilinus*, respectively, with a predicted average amino acid length of 150 aa. Within the candidate OBPs, using the SignalP v6.0 (Teufel et al., 2022) signal peptides have been identified in numbers: 22 out of 29 from *N. cubanus*, 29 out of 34 from *P. simplex*, and 17 out of 25 from *I. inquilinus*. Based on sequence analysis we identified, 3, 2 and 2 transcripts each, respectively, from *N. cubanus*, *P. simplex*, and *I. inquilinus* as Plus-C OBPs and 4, 3, 3, respectively, from the same as Minus-C OBPs. Adding OBP sequences from other insect orders we constructed a maximum likelihood phylogeny to classify the new candidates. Besides the newly identified sequences, we included OBP protein sequences from two other termite species, i.e., *Z. nevadensis* (Terrapon et al., 2014) and *C. secundus* (Harrison et al., 2018), as well as OBPs from representative species of Diptera, Coleoptera, Lepidoptera, and Hymenoptera (Große-Wilde et al., 2006; Vogt et al., 2015; Brand et al., 2018). Finally, we used OBPs of the basal hexapod *L. y-signata* as an outgroup (Missbach et al., 2015). The analysis allowed us to associate our candidates with four of the major OBP sub-groups: classic, Minus-C, Plus-C and ABP-II types (Figure 5). However, no 'Dimer-OBPs' or 'chemical-sense-related lipophilic-ligand-binding protein (CRLBP; Hekmat-Scafe et al., 2002) orthologs were found in termites. Additionally, we identified six Isoptera-specific expansions in the phylogeny. The 'Plus-C' subgroup contained two to three OBPs from each of our termite species was found at a basal position as well as at multiple clades in the phylogeny; classic and Minus-C OBP-subgroups each formed multiple clades in the phylogeny. Compared to other subgroups the most recently evolved Minus-C OBPs formed order-specific

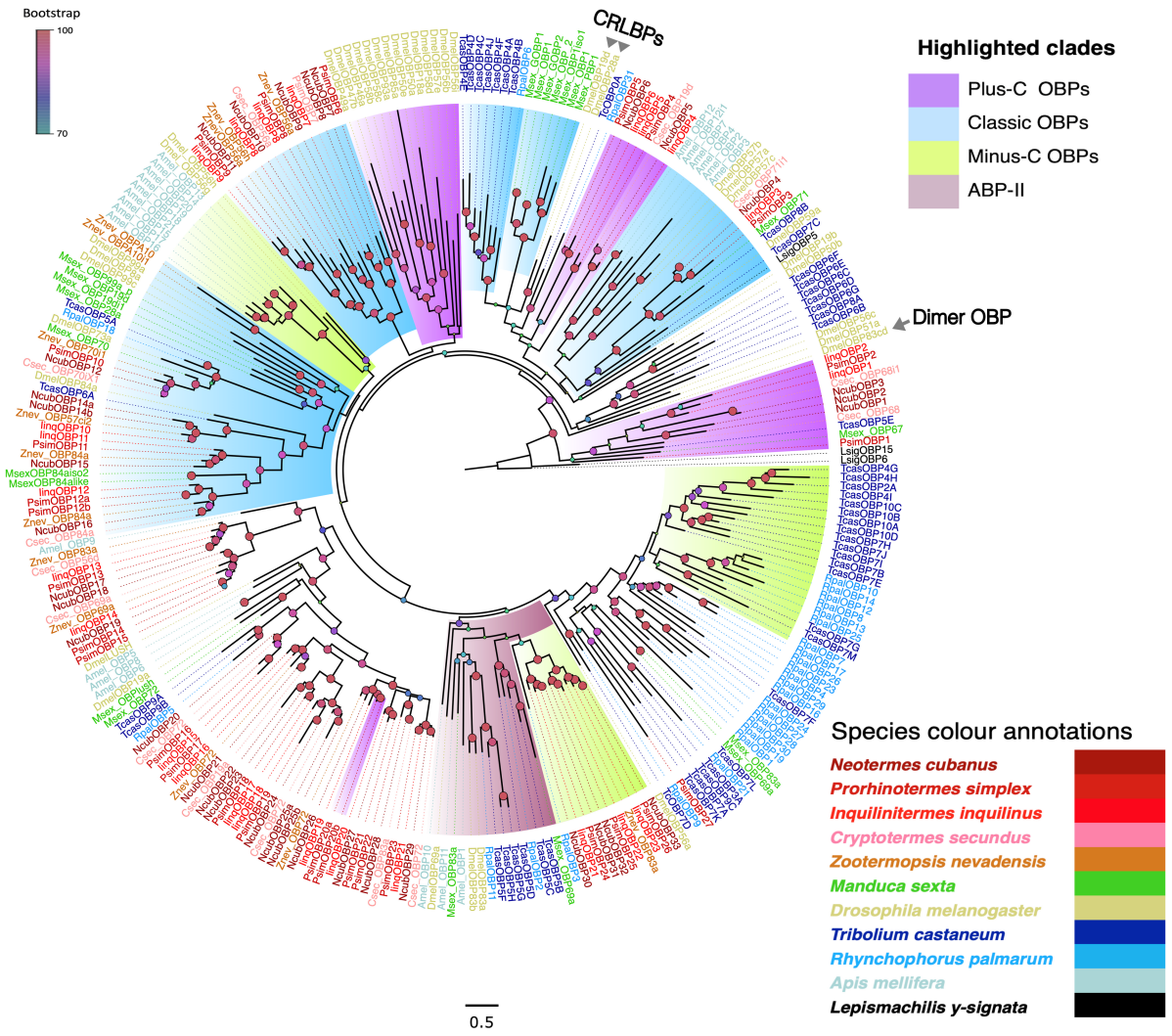

FIGURE 5
Maximum likelihood phylogeny of termite OBPs. The tree was constructed using LG+R as the best-fit amino acid substitution model and using *L. y-signata* OBPs as outgroup. The major OBP groups are highlighted as Classic OBPs (blue), Minus-C (green), Plus-C (violet) and ABP-II (maroon); due to low representation Dimer OBPs (DmelOBP83c) and CRLBPs (DmelOBP19D and DmelOBP28a) reported from *D. melanogaster* are not highlighted, but labelled. Sequence names are colored according to species and color codes are provided in the in-figure legend. The species included are *C. secundus* (Csec), *N. cubanus* (Ncub), *P. simplex* (Psim), *I. inquilinus* (Iinq), *Z. nevadensis* (Znev), *D. melanogaster* (Dmel), *T. castaneum* (Tcas), *R. palmarum* (Rpal), *M. sexta* (Msex), *Apis mellifera* (Amel) and *L. y-signata* (Lsig). The node colors indicate bootstrap support and the scale bar represents estimated amino acid substitutions per site.

expansions (Blattodea, Coleoptera and Lepidoptera). Termite OBPs also possessed orthologs in multiple Isoptera-specific expansions.

Next, we screened the transcriptomes for chemosensory proteins (CSPs), identifying 10, 6 and 9 CSPs from *N. cubanus*, *P. simplex*, and *I. inquilinus*, respectively. We further examined these proteins by reconstructing a maximum likelihood phylogeny, using CSPs reported for *R. speratus* (Mitaka et al., 2016), and reference CSPs from species in other insect orders, while using *D. pulex* CSPs as an outgroup (Figure 6). The phylogeny revealed species-specific CSP expansions in *D. melanogaster*, *T. castaneum*, and *B. mori*, but not in termites. There were two evolutionary patterns observed in CSPs, one a highly divergent clade of CSPs from a large number of species and a second one with mostly single orthologs from each species.

Discussion

Here, we present the analysis of the main chemosensory families (ORs, GRs, IRs, SNMPs, OBPs, and CSPs) in three species of termites from different phylogenetic lineages, encompassing basal and advanced clades. The number of ORs in the transcriptomes of *N. cubanus* (30), *P. simplex* (50), and *I. inquilinus* (28) was roughly similar to that reported from the genomes of *Z. nevadensis* (69) and *C. secundus* (42), and higher than

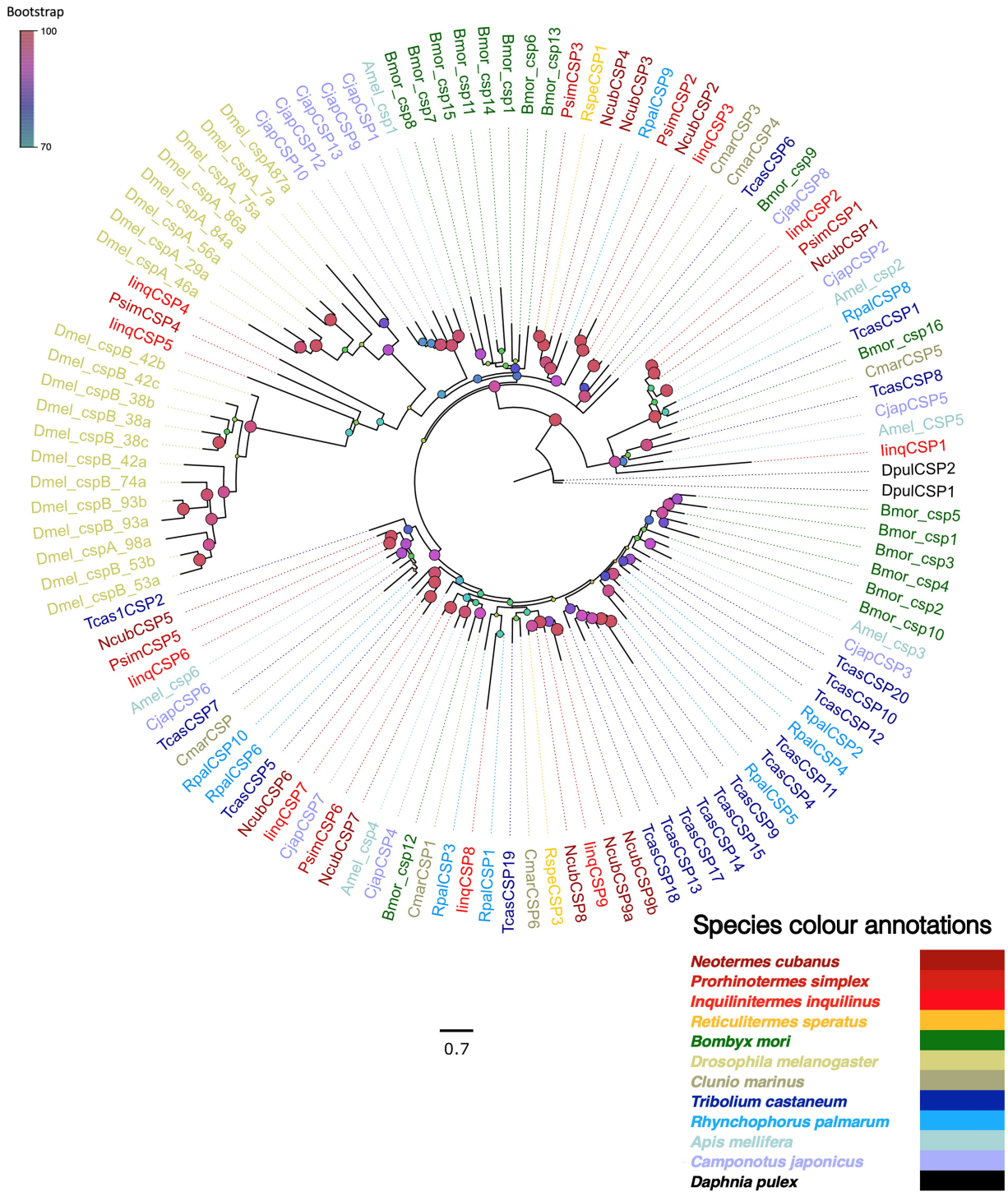

FIGURE 6
Maximum likelihood phylogeny of termite CSPs. The tree was constructed using LG+R as amino acid substitution model and rooted with *D. pulex* CSP sequences (Dpul, black) as outgroup. The newly identified termites CSPs were named as *N. cubanus* (Ncub), Psim: *P. simplex* (Psim), *I. inquilinus* (linq). The other species included were *R. speratus* (Rspe), *T. castaneum* (Tcas), *R. palmarum* (Rpal), *B. mori* (Bmor), *A. mellifera* (Amel), *D. melanogaster* (Dmel), *C. marinus* (Cmar), and *C. japonicus* (Cjap). Sequence names are colored according to species and color codes are provided in the in-figure legend. The node colors represent bootstrap support with 1,000 replicates and the scale bar represents estimated amino acid substitutions per site.

R. speratus (22; Terrapon et al., 2014; Mitaka et al., 2016; Harrison et al., 2018). However, the highest number of ORs among Blattodea was found in *B. germanica* (134 ORs), which may be partially explained by its large genome size, chromosomal translocations, and a higher rate of gene family expansions (Harrison et al., 2018). Ants, belonging among eusocial

Hymenoptera, also possess massive OR expansions, leading to ~350 ORs in *H. saltator* and *Camponotus floridanus* (Zhou et al., 2012). It has been hypothesized that these expansions are connected to their eusocial behavior (Zhou et al., 2012). However, high number of ORs has been reported also in non-eusocial Hymenoptera, such as *Nasonia vitripennis* (301, Robertson et al., 2010), suggesting that the expansion of ORs is an ancestral trait shared by Hymenoptera, which might potentially had facilitated the multiple independent evolutions of eusociality in hymenopteran insects. By contrast, the high OR repertoire reported recently in basal solitary apoid wasps phylogenetically positioned between ants and bees indicates that the OR repertoire in fact reduced during the evolution of eusocial apoids (Obiero et al., 2021).

Termites, despite being eusocial insects, exhibit numbers of OR genes comparable to non-eusocial insects (Mitaka and Akino, 2021). If an expansion of ORs preceded the emergence of eusociality in Hymenoptera, the same is clearly not true for Isoptera. The phylogenetic analysis revealed the highly conserved ORCo lineage and multiple Isoptera-specific OR expansions, which were analogous to the recent report in *C. secundus* based on the gene tree analysis (Harrison et al., 2018). Within these Isoptera-specific expansions, we found 1:1 orthologous relationship between the ORs of distinct termite species. This is rather unusual for the highly divergent OR family, indicating a high degree of OR conservation across termites.

The transcriptome screening performed in the three species of termites yielded 20, 25 and 26 GRs, respectively, from *N. cubanus*, *P. simplex*, and *I. inquilinus*. The phylogeny (Figure 2) also reveals the isopteran specific expansion of GRs in all three major sub clades: sugar, bitter and CO_2 receptors. The dendrogram clearly separates different GR sub-classes as taste and CO_2 and pheromone receptors. The basal clade includes *D. melanogaster* GR5a and Gr64a, which are tuned towards trehalose and sucrose, respectively, and who exhibit complementary functional profiles in *D. melanogaster* (Jiao et al., 2008). We have identified 6, 3 and 4 putative sugar receptors each, respectively, from *N. cubanus*, *P. simplex*, and *I. inquilinus*. The CO_2 receptor-containing clades was also largely expanded in termites. There were no orthologs found for Drosophila pheromone receptors GR32a and Gr68a. The fructose receptor (Gr43a) sub clade was located within the large bitter receptor clade, as previously reported in the *B. germanica* GR expansions (Robertson et al., 2018), indicating a conserved phylogenetic pattern across insect orders.

IRs, a subfamily of iGluRs, were found to be involved in detecting environmental as well as intracellular chemical signals (Benton et al., 2009; Croset et al., 2010; Ai et al., 2013). They were first identified in *D. melanogaster* and are well described in terms of the functional and evolutionary origins (Croset et al., 2010; Rytz et al., 2013). In contrast to other insect orders, IRs are numerous in Isoptera; in fact, the IR expansion in termites is considered to be analogous to OR expansions in Hymenoptera, signifying the importance of this protein family (Harrison et al., 2018; Robertson et al., 2018). A recent genome-based annotation in the cockroach *B. germanica* recovered 455 IRs, the highest number reported in insects. Nevertheless, nearly half of them were pseudogenes (Harrison et al., 2018; Robertson et al., 2018). While our findings of 98, 95 and 77 IRs from *N. cubanus*, *P. simplex*, and *I. inquilinus*, respectively, exceed the numbers identified in most other insect species, they fall in the range of the numbers reported from other termites (*Z. nevadensis*: 141; *C. secundus*: 135; Terrapon et al., 2014, Harrison et al., 2018). Although the olfactory perception of social signals in ants (Slone et al., 2017; Trible et al., 2017) and TFPs in termites (Gao et al., 2020) have been demonstrated as OR/ORCo dependent, it was also proposed that a parallel ionotropic receptor gene family expansion has favoured the evolution of colony communication in termites (Harrison et al., 2018). It seems likely that the total count of IR coding genes will be higher within the full genomes, but not substantially so. In contrast, only 12 IRs reported from *R. speratus* could be explained by the limited coverage of chemosensory genes in the whole-body transcriptome (Mitaka et al., 2016). It should be noted that genes with expression limited to one or a few tissues, like antennal IRs, will be underrepresented in a whole-body RNA pool, a reason we used antennal transcriptomes in our study. The expansion and positive selection in IRs have been reported recently in *Z. nevadensis* and *B. germanica* (Harrison et al., 2018). Rapid expansions in chemosensory receptor gene families provide functional divergence, crucial for adaption to different niches (Arguello et al., 2016). The caste and sex-biased expression of IRs reported in *Z. nevadensis* and *C. secundus* indicates the possible role of these genes in the pheromone communication (Harrison et al., 2018). The different subsets of iGluRs including IRs were added to the phylogenetic analysis. iGluRs exist across kingdoms, including plants, animals and prokaryotes (Croset et al., 2010; Rytz et al., 2013). We found orthologs of all major iGluR subfamilies in all three transcriptomes; our analysis revealed both antennal and divergent IRs. Similar to ORs, Isoptera-specific expansions were previously observed in termite IRs (Harrison et al., 2018). The antennal IRs are considered to be involved in olfaction, divergent IRs in gustation (Benton et al., 2009; Croset et al., 2010; Abuin et al., 2011; Prieto-Godino et al., 2017). As per the functional studies in *Drosophila*, the IR20a clade includes both taste and pheromone receptors (Koh et al., 2014). In our analysis this clade grouped with the divergent IRs. We identified five candidates in this clade, two each from *N. cubanus*, *P. simplex*, and one from *I. inquilinus*, and further research is required to confirm the role of these receptors.

SNMPs are broadly conserved CD36 (cluster of differentiation 36) family of transmembrane proteins in animals and are reported to be involved in the detection of lipid-derived pheromones in insects (Benton et al., 2007; Pregitzer et al., 2014). Among the two SNMP types reported in insects, SNMP1 was found to be expressed in both sensory neurons and supporting cells of insect pheromone-sensitive

sensilla, whereas SNMP2 was found only in the sensory supporting cells as reported in the moths *Heliothis virescens* and *Antheraea polyphemus* (Forstner et al., 2008). The recent structural studies indicate that SNMP1 might function as a co-receptor or acts as a tunnel to pass the signal molecules to the pheromone receptor (Gomez-Diaz et al., 2016). The number of SNMP1 proteins identified from our antennal transcriptomes was similar to the number reported from *C. secundus* (5) and higher than the one reported from *Z. nevadensis* (Terrapon et al., 2014; Harrison et al., 2018). Like termite ORs and IRs, SNMP1 showed 1:1 orthologous pattern among the five termite species compared. The higher number of SNMP1 proteins in termites could be correlated with the pheromone diversity in termites (Mitaka and Akino, 2021). In SNMP2 proteins, we found a single orthologous transcript in all five termite species compared. SNMP2, proteins are mainly found in the sensory neuron supporting cells and are proposed to be involved in pheromone clearance processes (Forstner et al., 2008).

OBPs and CSPs expressed in antennae and pheromone glands, respectively, are involved in both the reception and broadcast of the chemical message (Pelosi et al., 2018a). OBPs are highly abundant in the insect sensillar lymph and thus found abundantly in antennal transcriptomes (Venthur and Zhou, 2018). In Isoptera, OBPs and CSPs have been found to be differentially expressed among castes (Mitaka et al., 2016). The number of OBPs identified, i.e., 37, 35, 28 from *N. cubanus*, *P. simplex*, and *I. inquilinus*, respectively, are higher than the OBPs reported from other termites (*Z. nevadensis*: 19; *C. secundus*: 19; *R. speratus*: 9; Mitaka et al., 2016). Since OBPs are highly divergent in amino acid composition, using a basal hexapod *L. y-signata* (Missbach et al., 2015) as an outgroup helped in understanding OBP evolutionary pattern (Pelosi et al., 2005). All four major OBP sub-groups (classic, Minus-C, Plus-C and ABP-II types) have been identified based on the structure-based annotations reported earlier (Venthur et al., 2014). Unlike in Lepidoptera and other insect orders, Isopteran OBPs are understudied. However, we found two transcripts each in our transcriptomes with similarity to the well-studied protein *BmorPBP* from the moth *B. mori* (Lautenschlager et al., 2007). The 1:1 orthologous pattern observed in the other termite chemosensory genes continued in the case of OBPs. The number of CSPs identified was also higher in our transcriptomes as these were not annotated from the genomes of the other two termite species *Z. nevadensis* and *C. secundus* (Harrison et al., 2018).

Conclusion

Our research provides candidate genes of the major insect chemosensory gene families from three termite species belonging to three families of Isoptera of different phylogenetic positions, life histories and social complexities. We found comparatively large repertoires of chemosensory genes in all studied gene families as in other analysed termite species. The evolutionary analysis of termite chemosensory proteins revealed Isoptera-specific expansions with 1:1 orthologous pattern, indicating the existence of conserved olfactory functions. Our findings on basal eusocial insects will further enhance our understanding of the molecular underpinnings of eusociality.

Data availability statement

The datasets generated for this study can be found in the NCBI repository: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA885453/; accession numbers SRR21753787, SRR21753788, SRR21753789.

Ethics statement

Ethical review and approval was not required for the study of animals in accordance with the local legislation and institutional requirements.

Author contributions

OL and RH performed initial sequencing. Bioinformatic analysis was performed mainly by JJ, with assistance by SD, OL, MS, and EG-W. The study was conceived by RH and EG-W. Data interpretation and manuscript drafting was shared by all authors. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the grant No. 20-17194S from the Czech Science Foundation and by Institute of Organic Chemistry and Biochemistry, CAS (RVO: 61388963). EG-W was further supported by the EXTEMIT-K project financed by Operational Programme Research, Development and Education at the Czech University of Life Sciences, Prague, Czech Republic (CZ.02.1.01/0.0/0.0/15–003/0000433). RH was also supported by OP RDE as part of project grant no. CZ.02.1.01/0.0/0.0/16_019/0000803, "Advanced research supporting the forestry and wood-processing sector's adaptation to global change and the 4th industrial revolution."

Acknowledgments

We acknowledge Tereza Volštátová and Andrea Lepiešová of Czech University of Life Sciences Prague, Czech Republic for additional support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2022.1065947/full#supplementary-material

References

Abuin, L., Bargeton, B., Ulbrich, M. H., Isacoff, E. Y., Kellenberger, S., and Benton, R. (2011). Functional architecture of olfactory ionotropic glutamate receptors. *Neuron* 69, 44–60. doi: 10.1016/j.neuron.2010.11.042

Abuin, L., Prieto-Godino, L. L., Pan, H., Gutierrez, C., Huang, L., Jin, R., et al. (2019). *In vivo* assembly and trafficking of olfactory ionotropic receptors. *BMC Biol.* 17, 1–15. doi: 10.1186/s12915-019-0651-7

Ai, M., Blais, S., Park, J. Y., Min, S., Neubert, T. A., and Suh, G. S. B. (2013). Ionotropic glutamate receptors IR64a and IR8a form a functional odorant receptor complex *in vivo* in Drosophila. *J. Neurosci.* 33, 10741–10749. doi: 10.1523/JNEUROSCI.5419-12.2013

Andersson, M. N., Keeling, C. I., and Mitchell, R. F. (2019). Genomic content of chemosensory genes correlates with host range in wood-boring beetles (*Dendroctonus ponderosae*, *Agrilus planipennis*, and *Anoplophora glabripennis*). *BMC Genomics* 20, 690–617. doi: 10.1186/s12864-019-6054-x

Arguello, J. R., Cardoso-Moreira, M., Grenier, J. K., Gottipati, S., Clark, A. G., and Benton, R. (2016). Extensive local adaptation within the chemosensory system following *Drosophila melanogaster*'s global expansion. *Nat. Commun.* 7. doi: 10.1038/ncomms11855

Auer, T. O., Khallaf, M. A., Silbering, A. F., Zappia, G., Ellis, K., Álvarez-Ocaña, R., et al. (2020). Olfactory receptor and circuit evolution promote host specialization. *Nature* 579, 402–408. doi: 10.1038/s41586-020-2073-7

Bagnères, A.-G., and Hanus, R. (2015). "Communication and Social Regulation in Termites," in *Social Recognition in Invertebrates*. eds. L. Aquiloni and E. Tricarico (Cham: Springer International Publishing), 193–248.

Benton, R., Vannice, K. S., Gomez-Diaz, C., and Vosshall, L. B. (2009). Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. *Cells* 136, 149–162. doi: 10.1016/j.cell.2008.12.001

Benton, R., Vannice, K. S., and Vosshall, L. B. (2007). An essential role for a CD36-related receptor in pheromone detection in Drosophila. *Nature* 450, 289–293. doi: 10.1038/nature06328

Bohbot, J., Pitts, R. J., Kwon, H.-W., Rützler, M., Robertson, H. M., and Zwiebel, L. J. (2007). Molecular characterization of the *Aedes aegypti* odorant receptor gene family. *Insect Mol. Biol.* 16, 525–537. doi: 10.1111/j.1365-2583.2007.00748.x

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bordereau, C., and Pasteels, J. M. (2010). "Pheromones and chemical ecology of dispersal and foraging in termites," in *Biology of termites: A modern synthesis*. eds. D. E. Bignell, Y. Roisin and N. Lo (Dordrecht: Springer Netherlands), 279–320.

Brand, P., Robertson, H. M., Lin, W., Pothula, R., Klingeman, W. E., Jurat-Fuentes, J. L., et al. (2018). The origin of the odorant receptor gene family in insects. *elife* 7, 1–13. doi: 10.7554/eLife.38340

Bucek, A., Šobotník, J., He, S., Shi, M., McMahon, D. P., Holmes, E. C., et al. (2019). Evolution of termite symbiosis informed by transcriptome-based phylogenies. *Curr. Biol.* 29, 3728–3734.e4. doi: 10.1016/j.cub.2019.08.076

Butterwick, J. A., del Mármol, J., Kim, K. H., Kahlson, M. A., Rogow, J. A., Walz, T., et al. (2018). Cryo-EM structure of the insect olfactory receptor Orco. *Nature* 560, 447–452. doi: 10.1038/s41586-018-0420-8

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinform.* 10:421. doi: 10.1186/1471-2105-10-421

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348

Castillo, P., Le, N., and Sun, Q. (2021). Comparative antennal morphometry and sensilla organization in the reproductive and non-reproductive castes of the formosan subterranean termite. *Insects* 12, 576, 1–18. doi: 10.3390/insects12070576

Clyne, P. J., Warr, C. G., and Carlson, J. R. (2000). Candidate taste receptors in drosophila. *Science* 287, 1830–1834. doi: 10.1126/science.287.5459.1830

Clyne, P. J., Warr, C. G., Freeman, M. R., Lessing, D., Kim, J., and Carlson, J. R. (1999). A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in Drosophila. *Neuron* 22, 327–338. doi: 10.1016/S0896-6273(00)81093-4

Croset, V., Rytz, R., Cummins, S. F., Budd, A., Brawand, D., Kaessmann, H., et al. (2010). Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet.* 6:e1001064. doi: 10.1371/journal.pgen.1001064

Dahanukar, A., Foster, K., Goes van Naters, W. M., and Carlson, J. R. (2001). A gr receptor is required for response to the sugar trehalose in taste neurons of Drosophila. *Nat. Neurosci.* 4, 1182–1186. doi: 10.1038/nn765

Dahanukar, A., Lei, Y. T., Kwon, J. Y., and Carlson, J. R. (2007). Two gr genes underlie sugar reception in Drosophila. *Neuron* 56, 503–516. doi: 10.1016/j.neuron.2007.10.024

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2017). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 27, 1164–1165. doi: 10.1093/bioinformatics/btr088.ProtTest

De Fouchier, A., Walker, W. B., Montagné, N., Steiner, C., Binyameen, M., Schlyter, F., et al. (2017). Functional evolution of Lepidoptera olfactory receptors revealed by deorphanization of a moth repertoire. *Nat. Commun.* 8, 1–11. doi: 10.1038/ncomms15709

Delventhal, R., and Carlson, J. R. (2016). Bitter taste receptors confer diverse functions to neurons. *elife* 5, 1–23. doi: 10.7554/eLife.11181

Dippel, S., Kollmann, M., Oberhofer, G., Montino, A., Knoll, C., Krala, M., et al. (2016). Morphological and transcriptomic analysis of a beetle chemosensory system reveals a Gnathal olfactory center. *BMC Biol.* 14:90. doi: 10.1186/s12915-016-0304-z

Dolejšová, K., Křivánek, J., Šťáfková, J., Horáček, N., Havlíčková, J., Roy, V., et al. (2022). Identification of a queen primer pheromone in higher termites. *Commun. Biol.* 5, 1165–1111. doi: 10.1038/s42003-022-04163-5

Du, Y., Grodowitz, M. J., and Chen, J. (2019). Electrophysiological responses of eighteen species of insects to fire ant alarm pheromone. *Insects* 10. doi: 10.3390/insects10110403

Engsontia, P., Sangket, U., Chotigeat, W., and Satasook, C. (2014). Molecular evolution of the odorant and gustatory receptor genes in lepidopteran insects: implications for their adaptation and speciation. *J. Mol. Evol.* 79, 21–39. doi: 10.1007/s00239-014-9633-0

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* 17, 368–376. doi: 10.1007/BF01734359

Forstner, M., Gohl, T., Gondesen, I., Raming, K., Breer, H., and Krieger, J. (2008). Differential expression of SNMP-1 and SNMP-2 proteins in pheromone-sensitive hairs of moths. *Chem. Senses* 33, 291–299. doi: 10.1093/chemse/bjm087

Freeman, E. G., Wisotsky, Z., and Dahanukar, A. (2014). Detection of sweet tastants by a conserved group of insect gustatory receptors. *Proc. Natl. Acad. Sci. U. S. A.* 111, 1598–1603. doi: 10.1073/pnas.1311724111

Frickey, T., and Lupas, A. (2004). CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* 20, 3702–3704. doi: 10.1093/bioinformatics/bth444

Gao, Q., and Chess, A. (1999). Identification of candidate Drosophila olfactory receptors from genomic DNA sequence. *Genomics* 60, 31–39. doi: 10.1006/geno.1999.5894

Gao, Y., Huang, Q., and Xu, H. (2020). Silencing Orco impaired the ability to Perceive Trail pheromones and affected locomotion behavior in two termite species. *J. Econ. Entomol.* 113, 2941–2949. doi: 10.1093/jee/toaa248

Gomez-Diaz, C., Bargeton, B., Abuin, L., Bukar, N., Reina, J. H., Bartoi, T., et al. (2016). A CD36 ectodomain mediates insect pheromone detection via a putative tunnelling mechanism. *Nat. Commun.* 7. doi: 10.1038/ncomms11866

Gomez-Diaz, C., Martin, F., Garcia-Fernandez, J. M., and Alcorta, E. (2018). The two main olfactory receptor families in drosophila, ORs and IRs: a comparative approach. *Front. Cell. Neurosci.* 12:253. doi: 10.3389/fncel.2018.00253

Gonzalez, F., Johny, J., Walker, W. B., Guan, Q., Mfarrej, S., Jakše, J., et al. (2021). Antennal transcriptome sequencing and identification of candidate chemoreceptor proteins from an invasive pest, the American palm weevil, *Rhynchophorus palmarum*. *Sci. Rep.* 11:8334, 1–14. doi: 10.1038/s41598-021-87348-y

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Große-Wilde, E., Svatoš, A., and Krieger, J. (2006). A pheromone-binding protein mediates the bombykol-induced activation of a pheromone receptor in vitro. *Chem. Senses* 31, 547–555. doi: 10.1093/chemse/bjj059

Guo, W., Ren, D., Zhao, L., Jiang, F., Song, J., Wang, X., et al. (2018). Identification of odorant-binding proteins (OBPs) and functional analysis of phase-related OBPs in the migratory locust. *Front. Physiol.* 9:984. doi: 10.3389/fphys.2018.00984

Hansson, B. S., and Stensmyr, M. C. (2011). Evolution of insect olfaction. *Neuron* 72, 698–711. doi: 10.1016/j.neuron.2011.11.003

Harrison, M. C., Jongepier, E., Robertson, H. M., Arning, N., Bitard-Feildel, T., Chao, H., et al. (2018). Hemimetabolous genomes reveal molecular basis of termite eusociality. *Nat. Ecol. Evol.* 2, 557–566. doi: 10.1038/s41559-017-0459-1

Hekmat-Scafe, D. S., Scafe, C. R., McKinney, A. J., and Tanouye, M. A. (2002). Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. *Genome Res.* 12, 1357–1369. doi: 10.1101/gr.239402

Inward, D., Beccaloni, G., and Eggleton, P. (2007). Death of an order: a comprehensive molecular phylogenetic study confirms that termites are eusocial cockroaches. *Biol. Lett.* 3, 331–335. doi: 10.1098/rsbl.2007.0102

Isono, K., and Morita, H. (2010). Molecular and cellular designs of insect taste receptor system. *Front. Cell. Neurosci.* 4:20. doi: 10.3389/fncel.2010.00020

Jiao, Y., Moon, S. J., Wang, X., Ren, Q., and Montell, C. (2008). Gr64f is required in combination with other gustatory receptors for sugar detection in drosophila. *Curr. Biol.* 18, 1797–1801. doi: 10.1016/j.cub.2008.10.009

Jones, W. D., Cayirlioglu, P., Grunwald Kadow, I., and Vosshall, L. B. (2007). Two chemosensory receptors together mediate carbon dioxide detection in drosophila. *Nature* 445, 86–90. doi: 10.1038/nature05466

Katoh, K., Rozewicki, J., and Yamada, K. D. (2017). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108

Keesey, I. W., Zhang, J., Depetris-Chauvin, A., Obiero, G. F., Gupta, A., Gupta, N., et al. (2022). Functional olfactory evolution in *Drosophila suzukii* and the subgenus Sophophora. *iScience* 25:104212. doi: 10.1016/j.isci.2022.104212

Koh, T. W., He, Z., Gorur-Shandilya, S., Menuz, K., Larter, N. K., Stewart, S., et al. (2014). The drosophila IR20a clade of ionotropic receptors are candidate taste and pheromone receptors. *Neuron* 83, 850–865. doi: 10.1016/j.neuron.2014.07.012

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35, 4453–4455. doi: 10.1093/bioinformatics/btz305

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Lautenschlager, C., Leal, W. S., and Clardy, J. (2007). Bombyx mori pheromone-binding protein binding non-pheromone ligands: implications for pheromone recognition. *Structure* 15, 1148–1154. doi: 10.1016/j.str.2007.07.013

Leal, W. S. (2011). Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* 58, 373–391. doi: 10.1146/annurev-ento-120811-153635

Matsumura, F., Coppel, H. C., and Tai, A. (1968). Isolation and identification of Termite Trail-following pheromone. *Nature* 219, 963–964. doi: 10.1038/219963a0

Matsuura, K., Himuro, C., Yokoi, T., Yamamoto, Y., Vargo, E. L., and Keller, L. (2010). Identification of a pheromone regulating caste differentiation in termites. *Proc. Natl. Acad. Sci. U. S. A.* 107, 12963–12968. doi: 10.1073/pnas.1004675107

McBride, C. S., and Arguello, J. R. (2007). Five drosophila genomes reveal nonneutral evolution and the signature of host specialization in the chemoreceptor superfamily. *Genetics* 177, 1395–1416. doi: 10.1534/genetics.107.078683

Missbach, C., Dweck, H. K. M., Vogel, H., Vilcinskas, A., Stensmyr, M. C., Hansson, B. S., et al. (2014). Evolution of insect olfactory receptors. *elife* 3, 1–22. doi: 10.7554/eLife.02115

Missbach, C., Vogel, H., Hansson, B. S., and Große-Wilde, E. (2015). Identification of odorant binding proteins and chemosensory proteins in antennal transcriptomes of the jumping bristletail *Lepismachilis y-signata* and the firebrat *Thermobia domestica*: Evidence for an independent OBP-OR origin. *Chem. Senses* 40, 615–626. doi: 10.1093/chemse/bjv050

Mitaka, Y., and Akino, T. (2021). A review of termite pheromones: multifaceted, context-dependent, and rational chemical communications. *Front. Ecol. Evol.* 8:595614. doi: 10.3389/fevo.2020.595614

Mitaka, Y., Kobayashi, K., Mikheyev, A., Tin, M. M. Y., Watanabe, Y., and Matsuura, K. (2016). Caste-specific and sex-specific expression of chemoreceptor genes in a termite. *PLoS One* 11, 1–16. doi: 10.1371/journal.pone.0146125

Mitaka, Y., Mori, N., and Matsuura, K. (2017). Multi-functional roles of a soldier-specific volatile as a worker arrestant, primer pheromone and an antimicrobial agent in a termite. *Proc. Biol. Sci.* 284:20171134. doi: 10.1098/rspb.2017.1134

Montell, C. (2009). A taste of the drosophila gustatory receptors. *Curr. Opin. Neurobiol.* 19, 345–353. doi: 10.1016/j.conb.2009.07.001

Obiero, G. F., Pauli, T., Geuverink, E., Veenendaal, R., Niehuis, O., and Große-Wilde, E. (2021). Chemoreceptor diversity in Apoid wasps and its reduction during the evolution of the pollen-collecting lifestyle of bees (Hymenoptera: Apoidea). *Genome Biol. Evol.* 13, 1–18. doi: 10.1093/gbe/evaa269

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. *Nat. Commun.* 8, 297–210. doi: 10.1038/s41467-017-00099-1

Pĕalva-Arana, D. C., Lynch, M., and Robertson, H. M. (2009). The chemoreceptor genes of the waterflea Daphnia pulex: many grs but no Ors. *BMC Evol. Biol.* 9, 8–11. doi: 10.1186/1471-2148-9-79

Pelosi, P., Calvello, M., and Ban, L. (2005). Diversity of odorant-binding proteins and chemosensory proteins in insects. *Chem. Senses* 30, i291–i292. doi: 10.1093/chemse/bjh229

Pelosi, P., Iovinella, I., Zhu, J., Wang, G., and Dani, F. R. (2018a). Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. *Biol. Rev.* 93, 184–200. doi: 10.1111/brv.12339

Pelosi, P., Zhu, J., and Knoll, W. (2018b). Odorant-binding proteins as sensing elements for odour monitoring. *Sensors* 18. doi: 10.3390/s18103248

Pregitzer, P., Greschista, M., Breer, H., and Krieger, J. (2014). The sensory neurone membrane protein SNMP1 contributes to the sensitivity of a pheromone detection system. *Insect Mol. Biol.* 23, 733–742. doi: 10.1111/imb.12119

Prieto-Godino, L. L., Rytz, R., Cruchet, S., Bargeton, B., Abuin, L., Silbering, A. F., et al. (2017). Evolution of acid-sensing olfactory circuits in Drosophilids. *Neuron* 93, 661–676.e6. doi: 10.1016/j.neuron.2016.12.024

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Ramdya, P., and Benton, R. (2010). Evolving olfactory systems on the fly. *Trends Genet.* 26, 307–316. doi: 10.1016/j.tig.2010.04.004

Robertson, H. M. (2015). The insect chemoreceptor superfamily is ancient in animals. *Chem. Senses* 40, 609–614. doi: 10.1093/chemse/bjv046

Robertson, H. M., Baits, R. L., Walden, K. K. O., Wada-Katsumata, A., and Schal, C. (2018). Enormous expansion of the chemosensory gene repertoire in the omnivorous German cockroach *Blattella germanica*. *J. Exp. Zool. Part B Mol. Dev. Evol.* 330, 265–278. doi: 10.1002/jez.b.22797

Robertson, H. M., and Wanner, K. W. (2006). The chemoreceptor superfamily in the honey bee, *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family. *Genome Res.* 16, 1395–1403. doi: 10.1101/gr.5057506

Robertson, H. M., Gadau, J., and Wanner, K. W. (2010). The insect chemoreceptor superfamily of the parasitoid jewel wasp *Nasonia vitripennis*. *Insect Mol. Biol.* 19, 121–136. doi: 10.1111/j.1365-2583.2009.00979.x

Robertson, H. M., Warr, C. G., and Carlson, J. R. (2003). Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 14537–14542. doi: 10.1073/pnas.2335847100

Roisin, Y., and Korb, J. (2011). "Social organisation and the status of workers in termites," in *Biology of termites: A modern synthesis*. eds. D. E. Bignell, Y. Roisin and N. Lo (Dordrecht: Springer Netherlands), 133–164.

Rupf, T., and Roisin, Y. (2008). Coming out of the woods: do termites need a specialized worker caste to search for new food sources? *Naturwissenschaften* 95, 811–819. doi: 10.1007/s00114-008-0387-7

Rytz, R., Croset, V., and Benton, R. (2013). Ionotropic receptors (IRs): chemosensory ionotropic glutamate receptors in Drosophila and beyond. *Insect Biochem. Mol. Biol.* 43, 888–897. doi: 10.1016/j.ibmb.2013.02.007

Saina, M., Busengdal, H., Sinigaglia, C., Petrone, L., Oliveri, P., Rentzsch, F., et al. (2015). A cnidarian homologue of an insect gustatory receptor functions in developmental body patterning. *Nat. Commun.* 6:6243. doi: 10.1038/ncomms7243

Sánchez-Gracia, A., Vieira, F. G., and Rozas, J. (2009). Molecular evolution of the major chemosensory gene families in insects. *Heredity* 103, 208–216. doi: 10.1038/hdy.2009.55

Saran, R. K., Millar, J. G., and Rust, M. K. (2007). Role of (3Z,6Z,8E)-dodecatrien-1-ol in trail following, feeding, and mating behavior of *Reticulitermes hesperus*. *J. Chem. Ecol.* 33, 369–389. doi: 10.1007/s10886-006-9229-2

Scalzotto, M., Ng, R., Cruchet, S., Saina, M., Armida, J., Su, C.-Y., et al. (2022). Comparative transcriptomics of the Drosophila olfactory subsystems identifies a support cell-expressed Osiris protein required for pheromone sensing. *bioRxiv*, 2022.03.09.483212. doi: 10.1101/2022.03.09.483212

Scott, K., Brady, R., Cravchik, A., Morozov, P., Rzhetsky, A., Zuker, C., et al. (2001). A chemosensory gene family encoding candidate gustatory and olfactory receptors in Drosophila. *Cells* 104, 661–673. doi: 10.1016/S0092-8674(01)00263-X

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351

Slone, J. D., Pask, G. M., Ferguson, S. T., Millar, J. G., Berger, S. L., Reinberg, D., et al. (2017). Functional characterization of odorant receptors in the ponerine ant, *Harpegnathos saltator*. *Proc. Natl. Acad. Sci. U. S. A.* 114, 8586–8591. doi: 10.1073/pnas.1704647114

Song, L., and Florea, L. (2015). Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. *Gigascience* 4:48. doi: 10.1186/s13742-015-0089-y

Terrapon, N., Li, C., Robertson, H. M., Ji, L., Meng, X., Booth, W., et al. (2014). Molecular traces of alternative social organization in a termite genome. *Nat. Commun.* 5:3636, 1–12. doi: 10.1038/ncomms4636

Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gíslason, M. H., Pihl, S. I., Tsirigos, K. D., et al. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.* 40, 1023–1025. doi: 10.1038/s41587-021-01156-3

Thoma, M., Missbach, C., Jordan, M. D., Grosse-Wilde, E., Newcomb, R. D., and Hansson, B. S. (2019). Transcriptome surveys in silverfish suggest a multistep origin of the insect odorant receptor gene family. *Front. Ecol. Evol.* 7:281. doi: 10.3389/fevo.2019.00281

Trible, W., Olivos-Cisneros, L., McKenzie, S. K., Saragosti, J., Chang, N. C., Matthews, B. J., et al. (2017). Orco mutagenesis causes loss of antennal lobe glomeruli and impaired social behavior in ants. *Cells* 170, 727–735.e10. doi: 10.1016/j.cell.2017.07.001

Venthur, H., Mutis, A., Zhou, J. J., and Quiroz, A. (2014). Ligand binding and homology modelling of insect odorant-binding proteins. *Physiol. Entomol.* 39, 183–198. doi: 10.1111/phen.12066

Venthur, H., and Zhou, J.-J. (2018). Odorant receptors and odorant-binding proteins as insect pest control targets: a comparative analysis. *Front. Physiol.* 9:1163. doi: 10.3389/fphys.2018.01163

Vogt, R. G., Große-Wilde, E., and Zhou, J. J. (2015). The Lepidoptera odorant binding protein gene family: gene gain and loss within the GOBP/PBP complex of moths and butterflies. *Insect Biochem. Mol. Biol.* 62, 142–153. doi: 10.1016/j.ibmb.2015.03.003

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A., and Axel, R. (1999). A spatial map of olfactory receptor expression in the drosophila antenna. *Cells* 96, 725–736. doi: 10.1016/S0092-8674(00)80582-6

Vosshall, L. B., and Stocker, R. F. (2007). Molecular architecture of smell and taste in drosophila. *Annu. Rev. Neurosci.* 30, 505–533. doi: 10.1146/annurev.neuro.30.051606.094306

Yan, H., Jafari, S., Pask, G., Zhou, X., Reinberg, D., and Desplan, C. (2020). Evolution, developmental expression and function of odorant receptors in insects. *J. Exp. Biol.* 223:jeb208215, 1–11. doi: 10.1242/jeb.208215

Zhang, H. J., Anderson, A. R., Trowell, S. C., Luo, A. R., Xiang, Z. H., and Xia, Q. Y. (2011). Topological and functional characterization of an insect gustatory receptor. *PLoS One* 6, 1–10. doi: 10.1371/journal.pone.0024111

Zhao, Y. J., Li, G. C., Zhu, J. Y., and Liu, N. Y. (2020). Genome-based analysis reveals a novel SNMP group of the Coleoptera and chemosensory receptors in *Rhaphuma horsfieldi*. *Genomics* 112, 2713–2728. doi: 10.1016/j.ygeno.2020.03.005

Zhou, J. J., Kan, Y., Antoniw, J., Pickett, J. A., and Field, L. M. (2006). Genome and EST analyses and expression of a gene family with putative functions in insect chemoreception. *Chem. Senses* 31, 453–465. doi: 10.1093/chemse/bjj050

Zhou, X., Rokas, A., Berger, S. L., Liebig, J., Ray, A., and Zwiebel, L. J. (2015). Chemoreceptor evolution in Hymenoptera and its implications for the evolution of eusociality. *Genome Biol. Evol.* 7, 2407–2416. doi: 10.1093/gbe/evv149

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 Genet.* 8:e1002930. doi: 10.1371/journal.pgen.1002930