Received: 10 November 2020 | Revised: 22 March 2021 | Accepted: 26 March 2021

DOI: 10.1111/mec.15921

ORIGINAL ARTICLE

Smelling in the dark: Phylogenomic insights into the chemosensory system of a subterranean beetle

Pau Balart-García[1] | Alexandra Cieslak[1] | Paula Escuer[2] | Julio Rozas[2] | Ignacio Ribera[1] | Rosa Fernández[1]

[1]Institute of Evolutionary Biology (CSIC - Universitat Pompeu Fabra), Barcelona, Spain

[2]Department of Genetics, Microbiology and Statistics, Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Spain

Correspondence
Pau Balart-García and Rosa Fernández, Institute of Evolutionary Biology (CSIC - Universitat Pompeu Fabra), Barcelona, Spain.
Email: pau.balart@ibe.upf-csic.es (PB-G); rosa.fernandez@ibe.upf-csic.es (RF)

Funding information
Ministerio de Economía y Competitividad and the Ministerio de Ciencia of Spain, Grant/Award Number: CGL2016-76705-P, PID2019-108824GA-I00, CGL2016-75255, PID2019-103947GB, BES-2017-081050, BES-2017-081740 and RyC2017-22492; Marie Sklodowska-Curie, Grant/Award Number: 747607

Abstract

The chemosensory system has experienced relevant changes in subterranean animals, facilitating the perception of specific chemical signals critical to survival in their particular environment. However, the genomic basis of chemoreception in cave-dwelling fauna has been largely unexplored. We generated de novo transcriptomes for antennae and body samples of the troglobitic beetle *Speonomus longicornis* (whose characters suggest an extreme adaptation to a deep subterranean environment) in order to investigate the evolutionary origin and diversification of the chemosensory gene repertoire across coleopterans through a phylogenomic approach. Our results suggested a diminished diversity of odourant and gustatory gene repertoires compared to polyphagous beetles that inhabit surface habitats. Moreover, *S. longicornis* showed a large diversity of odourant-binding proteins, suggesting an important role of these proteins in capturing airborne chemical cues. We identified a gene duplication of the ionotropic coreceptor IR25a, a highly conserved single-copy gene in protostomes involved in thermal and humidity sensing. In addition, no homologous genes to sugar receptors or the ionotropic receptor IR41a were detected. Our findings suggest that the chemosensory gene repertoire of this cave beetle may result from adaptation to the highly specific ecological niche it occupies, and that gene duplication and loss may have played an important role in the evolution of gene families involved in chemoreception. Altogether, our results shed light on the genomic basis of chemoreception in a cave-dwelling invertebrate and pave the road towards understanding the genomic underpinnings of adaptation to the subterranean lifestyle at a deeper level.

KEYWORDS
chemosensory proteins, Coleoptera, transcriptomics, troglobitic fauna

1 | INTRODUCTION

Major lifestyle transitions in insects, such as the conquest of terrestrial habitats, flight or host-plant interactions, are often accompanied by dramatic shifts in their sensory systems (Almudi et al., 2020; Anholt, 2020; Missbach et al., 2015; Vieira & Rozas, 2011; Wang, Pentzold, et al., 2018). Subterranean specialization has also offered opportunities for evolutionary innovation in the way animals interact with this particular environment (Cartwright et al., 2017). While adapting to subterranean niches, different species,

Ignacio Ribera and Rosa Fernández are Senior authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2021 The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ranging from fishes to insects, have evolved highly convergent alternatives to live in perpetual darkness in habitats exhibiting specific biotic and abiotic factors (i.e., limited nutrient sources, constant temperature and humidity). Evolutionary regressions (e.g., loss of eyes and pigmentation), elaborated elements (e.g., hypertrophy of extra-optic sensory structures) and other physiological changes (e.g., modified life cycles) have been reported as possible adaptations for many obligate subterranean fauna (Culver & Pipan, 2019). Likewise, it is conceivable that the subterranean selective pressures have driven adaptive shifts in other sensory systems, including the chemosensory systems of subterranean animals. For instance, some studies on cavefish pointed out an enhancement of chemosensory systems from a morphological point of view (i.e., visible differences in taste buds and olfactory neural bulbs) when compared to surface populations (Parzefall, 2001; Yamamoto et al., 2009; Yang et al., 2016). In subterranean arthropods, elongation of antennae and body appendages have also been attributed to enhanced sensory capabilities (Turk et al., 1996). Nevertheless, the evolution of the chemosensory system in subterranean fauna from a molecular perspective remains widely unexplored.

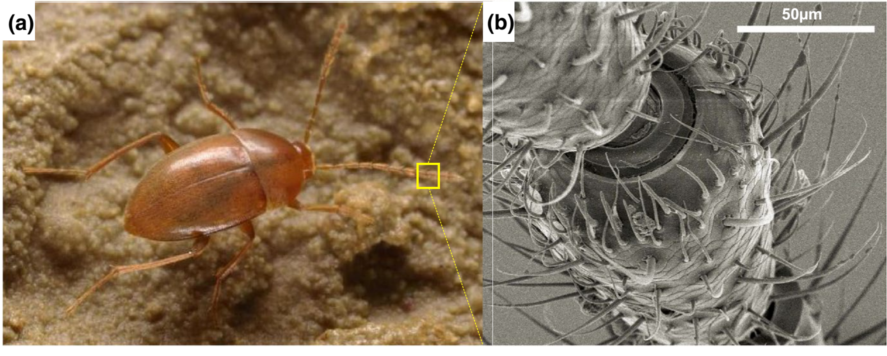
Environmental chemical signals are enormously diverse in nature. Animals have developed a wide diversity of mechanisms to perceive and interpret specific cues essential to their evolutionary success (Nei et al., 2008). In insects, these chemicals include palatable nutrient or repellent odours and tastes, pheromones, warning signals of predators and those indicating optimal substrates for oviposition, and various others (Joseph & Carlson, 2015). The chemosensory system in insects is distributed morphologically at the interface between the environment and the dendrites of the peripheral sensory neurons, where different chemosensory proteins act in parallel for the signal transduction to the brain centers in which the information is processed (Dippel et al., 2016; Joseph & Carlson, 2015). To capture this complex information, insects use three large and divergent families of transmembrane chemoreceptor proteins: gustatory receptors (GRs), odourant receptors (ORs) and ionotropic receptors (IRs; Benton et al., 2009; Clyne et al., 1999; Gao & Chess, 1999; Sánchez-Gracia et al., 2009; Vosshall et al., 1999). GRs, which detect nonvolatile compounds, probably represent the oldest chemosensory receptors (Eyun et al., 2017), being distributed in several taste organs along the entire body including mouth pieces, legs, wing margins and other specialized structures such as vaginal plate sensilla in abdomens of female flies (Stocker, 1994). Airborne chemical particles are perceived in the head appendages by the ORs, an insect-specific chemoreception gene family thought to have originated from the GR gene family (Robertson, 2019; Robertson et al., 2003; Thoma et al., 2019). ORs work with the functionally universal odourant receptor coreceptor (ORCO), which is highly conserved in winged insects (i.e., Paleoptera and Neoptera; Brand et al., 2018). Moreover, IRs derived from the ionotropic glutamate receptor gene (IGluRs) superfamily in protostomes (Benton, 2015; Vosshall & Stocker, 2007) and mediate responses to many organic acids and amines, including pheromones and nutrient odours (Benton et al., 2009). In insects there are other gene families that also participate in chemosensory functions, such as the sensory neuron membrane proteins (SNMPs; Grimaldi & Engel, 2005; Missbach et al., 2014; Nichols & Vogt, 2008). The odourant-binding proteins (OBPs) and chemosensory proteins (CSPs) also play a key role for chemoreception in terrestrial insects, besides other physiological roles. The stable and compact structure of OBPs and CSPs make them versatile soluble proteins relevant for signal transduction of small hydrophobic compounds such as pheromones and odourants (Pelosi et al., 2014, 2018; Roys, 1954; Stürckow, 1970).

As in other large gene families encoding ecologically relevant proteins, constant birth-and-death dynamics may play an important role in their evolution in arthropods (Nei & Rooney, 2005; Sánchez-Gracia et al., 2009; Vieira et al., 2007). A general positive correlation has been observed when comparing the chemosensory gene diversity across species and the complexity of chemical signal in the ecological niche they occupy; several studies reported contrasting patterns of gene family expansions and gene losses when exploring the chemosensory gene repertoires in extreme specialist and generalist species, with the latter usually exhibiting larger expansions of gene families involved in chemoreception (Andersson et al., 2019; Kirkness et al., 2010; Li et al., 2018; McBride, 2007; McBride & Arguello, 2007; Ngoc et al., 2016). However, the evolution of the chemosensory gene families in subterranean species are still largely unexplored, hampering our understanding on how these animals perceive their particular environment.

Cave beetles represent ideal models to shed light on the genomic basis of chemoreception in subterranean environments. The Leptodirini tribe is a speciose lineage of scavenger beetles that represents one of the most impressive radiations of subterranean organisms. Several lineages within Leptodirini (estimated to have colonized subterranean habitats ca. 33 Ma (Ribera et al., 2010)) acquired morphological and physiological traits typically associated with troglobitic adaptations. Their modifications include complete lack of eyes and optic lobes, depigmentation, membranous wings, elongation of antennae and legs (Deleurance, 1963; Jeannel, 1924; Luo et al., 2019) and loss of thermal acclimation capacity (Pallarés et al., 2020; Rizzo et al., 2015). They also exhibit modified life cycles as a key innovation for their subterranean specialization (Cieslak et al., 2014; Delay, 1978). One of the highly modified species of the Leptodirini tribe is *Speonomus longicornis* Saulcy, 1872 (Coleoptera, Polyphaga, Leiodidae; Figure 1a). This obligate cave-dwelling beetle is completely blind, depigmented, possesses enlarged antennae (Jeannel, 1924) with a high sensilla density (Figure 1b) and it has a contracted life cycle, comprising a single larval-instar during its development in which the larvae remain practically quiescent like the pupal stage (Glaçon, 1953). The troglobitic characters of this species suggest an extreme adaptation to the deep subterranean environment.

This study aims to characterize the chemosensory gene repertoire of *S. longicornis*. The aims of this project were (i) to pinpoint genes putatively involved in chemoreception in the cave beetle *S. longicornis* through a transcriptomic approach, and (ii) to explore

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



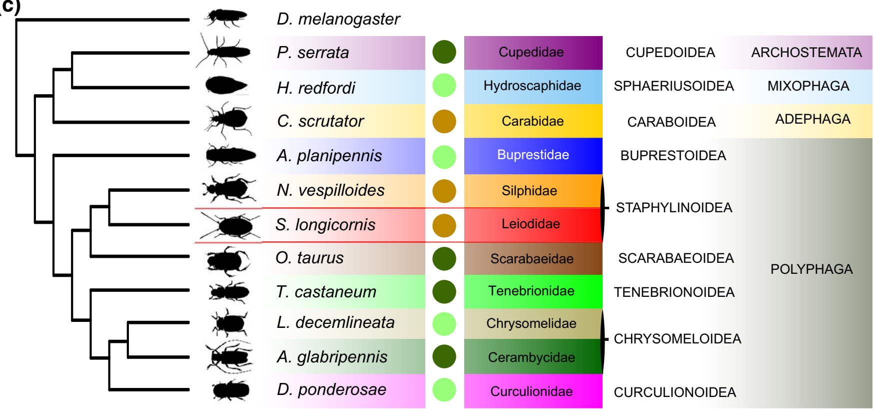


FIGURE 1 *Speonomus longicornis* and its phylogenetic position within Coleoptera. (a) *Speonomus longicornis*. (b) Scanning electron microscopy image of the antennal sensilla of *S. longicornis* (voucher IBE-AI531). (c) Simplified phylogeny showing the relationships of the studied species, adapted from McKenna et al. (2019). Coloured circles illustrate the dietary habits of the species: darkgreen corresponds to polyphagous herbivores, lightgreen to oligophagous herbivores and brown to nonphytophagous species [Colour figure can be viewed at wileyonlinelibrary.com]

how such genes evolved in the broader phylogenetic context of beetle and insect evolution. Our study therefore aims to provide the first characterization of the chemosensory gene repertoire of an obligate cave-dwelling species.

2 | MATERIALS AND METHODS

2.1 | Sample collection and preservation

Thirty specimens of *Speonomus longicornis* were collected in 2016 at the type locality: Grotte de Portel cave, in the Plantaurel massif at the French region of Ariège (43°01′51″N, 1°32′22″E). All specimens were manually captured and kept alive inside a thermo-box during the stay at the cave. Once sampling was completed, all individuals were placed in an 8 ml tube and flash-frozen in liquid nitrogen at the cave entrance in order to prevent stress-related alterations in gene expression levels and to minimize RNA degradation during transportation to the laboratory, where the samples were stored at −80°C until RNA extraction.

2.2 | RNA extraction

All steps were performed in cold and RNase free conditions. Several specimens were pooled in each sample prior to RNA extraction in order to obtain sufficient tissue for an efficient extraction. We did not examine the sex of the specimens to minimize the manipulation in order to avoid RNA degradation. Nevertheless, no significant sexual dimorphism has previously been found in the chemosensory system of other coleopterans (Dippel et al., 2016; Wu et al., 2016).

The specimens were split into three groups of 10 individuals each, representing biological replicates. Since chemosensory structures are mainly concentrated in the antennae (see Section 1), they were dissected from each specimen. Therefore, the experimental design included three biological replicates representing two conditions: antennae and the rest of the body.

The isolation of total RNA was performed by phenol/chloroform extraction with a lysis through guanidinium thiocyanate buffer following the protocol of Sambrook et al. (1989) with minor modifications (i.e., not using 2-mercaptoethanol). A first quality check was done by size separation in a 1% TBE agarose gel chromatography. Total RNA yield was quantified by an RNA assay in a Qubit fluorometer (Life Technologies).

2.3 | cDNA library construction and next-generation sequencing

For the antennae samples, a low-input RNA sequencing protocol was followed. mRNA sequencing libraries were prepared following the SMARTseq2 protocol (Picelli et al., 2013) with some modifications. Briefly, RNA was quantified using the Qubit RNA HS Assay Kit (Thermo Fisher Scientific). Reverse transcription with an input material of 2 ng was performed using SuperScript II (Invitrogen) in the presence of oligo-dT30VN (1 µM; 5′-AAGCAGTGGTATCAACGCAGAGTACT30VN-3′), template-switching oligonucleotides (1 µM) and betaine (1 M). The cDNA was amplified using

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the KAPA Hifi Hotstart ReadyMix (Roche), 100 nM ISPCR primer (5′-AAGCAGTGGTATCAACGCAGAGT-3′) and 15 cycles of amplification. Following purification with Agencourt Ampure XP beads (1:1 ratio; Beckmann Coulter), product size distribution and quantity were assessed on a Bioanalyzer High Sensitivity DNA Kit (Agilent). The amplified cDNA (200 ng) was fragmented for 10 min at 55°C using Nextera XT (Illumina) and amplified for 12 cycles with indexed Nextera PCR primers. The library was purified twice with Agencourt Ampure XP beads (0.8:1 ratio) and quantified on a Bioanalyzer using a high sensitivity DNA kit.

For the samples containing the rest of the body, total RNA was assayed for quantity and quality using the Qubit RNA BR Assay kit (Thermo Fisher Scientific) and RNA 6000 Nano Assay on a Bioanalyser 2100 (Agilent). The RNASeq libraries were prepared from total RNA using the KAPA Stranded mRNA-Seq Kit for Illumina (Roche) with minor modifications. Briefly, after poly-A based mRNA enrichment from 500 ng of total RNA, the mRNA was fragmented. The second strand cDNA synthesis was performed in the presence of dUTP to achieve strand specificity. The blunt-ended double stranded cDNA was 3′ adenylated and Illumina single indexed adapters (Illumina) were ligated. The ligation product was enriched with 15 PCR cycles and the final library was validated on an Agilent 2100 Bioanalyser with the DNA 7500 assay.

The libraries were sequenced on an Illumina HiSeq 2500 platform in paired-end mode with a read length of 2 × 76 bp. Image analysis, base calling and quality scoring of the run were processed using the manufacturer's software REAL TIME ANALYSIS (RTA 1.18.66.3) and followed by generation of FASTQ sequence files by CASAVA. cDNA libraries and mRNA sequencing were performed at the National Center of Genomic Analyses (CNAG).

2.4 | Sequence processing, decontamination and de novo assembly

Raw reads for all samples were downloaded in FASTQ format. The quality of the raw reads was assessed and visualized using FASTQC version 0.11.8 (www.bioinformatics.babraham.ac.uk). For each data set, remaining Illumina adaptors were removed and low-quality bases were trimmed off according to a threshold average quality score of 30 based on a Phred scale with TRIMMOMATIC version 0.38 (Bolger et al., 2014). Filtered paired-end reads were validated through a FASTQC visualization.

A reference de novo transcriptome assembly was constructed with TRINITY version 2.8.4, using paired read files and default parameters, including all replicates and conditions (Grabherr et al., 2011; Haas et al., 2013). BLOBTOOLS version 1.1.1 (Laetsch & Blaxter, 2017) was used to detect putative contamination from the assembled transcriptome. Transcripts were annotated using BLAST+ version 2.4.0 against the nonredundant (nr) database from NCBI with an expect value (*E*-value) cutoff of $1e^{-10}$ and reads were mapped to the reference transcriptome with BOWTIE2 version 2.3.5.1 (Langmead & Salzberg, 2012). Putative contaminants included transcripts with significant hits to viruses, fungi, bacteria or chordates, accounting for a total of 7.8% of the mapped sequences (see also Figure S1).

2.5 | Inference of candidate coding regions and transcriptome completeness assessment

To check completeness of the reference transcriptome, we searched for single copy universal genes in insects through benchmarking universal single-copy orthologues (BUSCO version 4.1: Simão et al., 2015) using the insecta database (insecta_odb10) and three different data sets as a query: (i) the assembled transcripts in TRINITY, (ii) the total predicted open reading frames (ORFs), and (iii) the longest isoform per ORF. The assembly was processed in TRANSDECODER version 5.4.0 to identify candidate ORFs within the transcripts using the universal genetic code (Haas et al., 2013). Only the longest ORFs (i.e., with a minimum length of 100 amino acids) of each transcript were retained as final candidate coding regions for further analyses.

2.6 | Chemosensory gene repertoire characterization

BITACORA version 1.0.0 (Vizueta et al., 2020) was used to curate annotations during the sequence similarity searches of the chemosensory gene families of interest. Curated protein databases containing chemoreceptor genes (ORs, GRs, IRs, SNMPs, OBPs and CSPs) of several arthropods were used in the BITACORA searches (Vizueta et al., 2016). An additional database was used for the ORs annotation, containing Coleoptera ORs based on the data sets from Mitchell et al. (2019). The "protein mode" pipeline of BITACORA was used to annotate all the predicted ORFs of the transcriptome (see previous section), combining BLAST and HMMER searches. All the predicted coding regions were used for BITACORA searches, retrieving a multifasta file for each of the chemosensory families. Dubious annotations were manually inspected (i.e., some ORFs received significant hits for both ORs and GRs, which were post-validated through Pfam searches). Results were filtered with customized Python scripts using BIOPYTHON version 1.76 SEQIO package (https://biopython.org [Cock et al., 2009]); in order to obtain final candidates, which were represented by the longest isoform per gene and thus achieving unique gene annotations.

2.7 | Expression levels quantification and differential gene expression analysis

SALMON version 0.10.2 (Patro et al., 2017) was used for indexing and quantification of transcript expression. Expression estimated counts were transformed into an expression matrix using a Perl script included in the TRINITY software (abundance_estimates_to_matrix.pl), which implements the trimmed mean of *M*-values normalization method (TMM). We also examined the data to ensure that the

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

biological replicates were well correlated using the "PtR" module included in TRINITY version 2.8.4 analysis toolkit, generating a variety of plots that allowed us to visually inspect the presence of strong outliers or batch effects that can affect the differential expression analysis. Differential expression analysis was conducted in the Bioconductor EDGER package (Robinson et al., 2010; Robinson & Oshlack, 2010). The Benjamini-Hochberg method was applied to control the false discovery rate (FDR; Benjamini & Hochberg, 1995). The significance value for multiple comparisons was adjusted to 0.001 FDR threshold cutoff and a four-fold change. Differentially expressed genes (up- and downregulated) in antennae and the rest of the body were plotted in heatmaps using the R scripts provided in the TRINITY software. The expression matrix was also interrogated in order to detect exclusively expressed genes in antennae (defined as genes that showed positive expression values in the three replicates of antennae and with expression values lower than 0.001 TMM in the rest of the body).

2.8 | Transcriptome characterization, gene ontology enrichment and visualization

The peptide predictions, including all isoforms, were used as input for EGGNOG-MAPPER version 4.5.1 (Huerta-Cepas et al., 2017), retrieving gene ontology (GO) terms for all the annotated transcripts. The GO annotations were subsequently filtered to discard those corresponding to nonanimal taxa (i.e., viruses, bacteria, fungi, plants, 10.86% of total GO annotations) and to eliminate the redundancy provided by the isoforms. All GO terms for each unique gene were retained and used in the subsequent analyses. GO enrichment analysis was performed using a two-tailed Fisher's test in FATIGO software (Al-Shahrour et al., 2007) to detect significant overrepresentation of the GO terms in the pairwise comparisons between the upregulated genes in antennae and the rest of the body, adjusting the *p*-value to .05.

GO enrichment analyses were visualized in the REVIGO web server (Supek et al., 2011), plotting the results in a "TREEMAP" graph using R, where the size of the rectangles is proportional to the enrichment *p*-value (abs_log10_pvalue) of the overrepresented GO terms.

2.9 | Phylogenetic inferences for the candidate chemosensory genes

Complete and partial annotated genes for *S. longicornis* (referred to as *Slon* in the figures) were included in the phylogenetic inferences in order to interrogate their phylogenetic relationships with chemosensory genes of other species, all of them based on genomic data. With this approach, the aim is to infer diversity patterns of the chemosensory repertoire of *S. longicornis* and to characterize each gene family more specifically, indicating with a higher confidence the putative function of these genes compared to analysis merely based on homology. Individual phylogenies for each chemosensory gene family as annotated by BITACORA (see above) were inferred using the following pipeline. Amino acid sequences were aligned using PASTA software version 1.7.8 (Mirarab et al., 2015). Poorly aligned regions were trimmed using TRIMAL version 1.2 (Capella-Gutiérrez et al., 2009) with the "-automated1" flag. Maximum likelihood phylogenetic inference was inferred with IQ-TREE version 2.0.4 (Nguyen et al., 2015). The mixture model LG + C20 + F + G was used with the site-specific posterior mean frequency model (PMSF; Wang et al., 2018) and the ultrafast bootstrap option (Hoang et al., 2018). A guide tree was inferred with FASTTREE2 under the LG model (Price et al., 2010). Results were visualized using the ITOL web interface (Letunic & Bork, 2019).

The ORs phylogeny included the coleopteran ORs obtained from Mitchell et al. (2019). These species (Figure 1c) included a range of ecological strategies. Considering the degree of dietary specialization, the phytophagous specialists (i.e., species feeding on a small number of plants or algae species usually belonging to the same botanical family, thus considered as oligophagous [Schoonhoven et al., 2005]) were represented by the aquatic beetle *Hydroscapha redfordi* (Myxophaga, Hydroscaphidae), the ash borer *Agrilus planipennis* (Polyphaga, Buprestidae), the Colorado potato beetle *Leptinotarsa decemlineata* (Polyphaga, Chrysomelidae) and the mountain pine beetle *Dendroctonus ponderosae* (Polyphaga, Curculionidae). By contrast, the phytophagous generalists (i.e., species feeding on several plants belonging to different families, thus considered as polyphagous) included the red flour beetle *Tribolium castaneum* (Polyphaga, Tenebrionidae), the wood borer *Anoplophora glabripennis* (Polyphaga, Cerambycidae), the reticulated beetle *Priacma serrata* (Archostemata, Cupedidae) and the dung beetle *Onthophagus taurus* (Polyphaga, Scarabaeidae). Moreover, the species set include two nonphytophagous beetles, the insectivorous *Calosoma scrutator* (Adephaga, Carabidae) and the burying beetle *Nicrophorus vespilloides* (Polyphaga, Silphidae) that feeds on vertebrate carrion (Vogel et al., 2017).

For the ORCO phylogeny, some additional ORCO sequences of coleopteran species and other taxa as outgroups (*Apis mellifera* and *Drosophila melanogaster*) were also included (see species and GenBank accessions at Table S1).

3 | RESULTS

3.1 | A high quality de novo transcriptome for *Speonomus longicornis* facilitates the annotation of its chemosensory gene repertoire

Since no reference genome is available for the focus species, a deeply sequenced de novo assembly transcriptome was constructed combining the paired-end reads from the six libraries (~411 million reads; ~356 million after trimming), obtaining a total of 245,131 transcripts (Figure S1). These transcripts include 177,711 unique predicted "genes" by TRINITY and 74,273 candidate ORFs (including all genes and isoforms). When filtering by the longest isoform per gene (which could be considered as a "proxy" for the total number of genes in the genome), we obtained a total of

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

20,956 ORFs. BUSCO analysis indicated a high completeness for the assembled transcriptome, with 97% of complete BUSCO genes compared to the insecta database (Figure S1), indicating that a mostly complete reference gene set was recovered and hence it was of enough quality to explore gene family evolution. These results should be interpreted with caution as they are based on transcriptomic data instead of high-quality genomes, and further reference-level genomic analyses may help clarify the evolutionary dynamics of chemosensory gene families with higher precision. However, this approach has been successfully applied in other studies with non-model organisms through the combination of genomic and high quality transcriptomic data (e.g., Fernández & Gabaldón, 2020; Vizueta, Escuer, et al., 2020). Further details about sequencing, assembly statistics, the completeness assessment and the putative contamination results are summarized in Figure S1.

3.2 | Differential gene expression analysis reveals chemosensory genes upregulated in the antennae

BITACORA searches identified a total of 205 chemosensory gene candidates for *S. longicornis* (Table 1, see also Table S2). The expression level distribution obtained in the transcript quantification steps was assessed in order to identify possible biases when comparing replicates and conditions (Figure S2). Results indicated that replicates are more similar to each other than between the different conditions. A total of 18,160 clusters of transcripts (reported as clusters of transcripts or "genes" by TRINITY, referred to as TRINITY genes hereafter) were detected as differentially expressed in antennae and body, with 8,949 TRINITY genes upregulated in antennae (Figure S3; Table S3). Out of the 205 candidate chemosensory genes as detected by BITACORA, 78 were detected as differentially expressed. From those, 49 genes were overexpressed in antennae (18 ORs, 17 OBPs, five GRs, five IRs/IGluRs, two SNMPs/CD36s and two CSPs; Table 1; Figure 2a). Furthermore, seven ORs were identified as exclusively expressed in antennae, including two additional genes not recovered as differentially expressed due to the disparity of the expression values between antennae replicates (Table S4).

3.3 | Gene ontology enrichment reveals upregulated chemosensory specificity in antennae

Out of the 74,273 predicted ORFs (including isoforms), 59.4% were annotated through EGGNOG-MAPPER, and from those only 23,555 of these annotations yielded associated GO terms, representing 31.7% of the total queried sequences from the assembled transcriptome. After filtering annotations from nonanimal taxa (including viruses, bacteria and fungi) 89.14% of the annotations were retained. Figure 2b depicts enriched GO terms for upregulated genes in antennae and in the rest of the body. In antennae, "sensation and perception of chemical stimulus" represent the most enriched category within the biological processes analysed and, in a minor proportion, some categories related to cilium activity. "Mechanosensory activity" terms are also overrepresented but in a minor proportion. More than half of the cellular component GO terms enriched in antennae correspond to "dendritic structures", and a high proportion of the overrepresented terms correspond to "extracellular and membrane structures". Regarding the molecular function category in antennae, "odourant-binding" and "odourant reception" terms occupy a large proportion of the enriched functions followed by other "binding and signal transduction" terms.

TABLE 1 Annotated chemosensory genes of *S. longicornis*

	ORs	GRs	IRs/IGluRs	SNMPs/CD36s	OBPs	CSPs
Total	50	36	53	20	39	7
Antennae	18	5	5	2	17	2

The number of overexpressed genes in antennae are indicated.

3.4 | Phylogenetic interrelationships of Coleoptera ORs and ORCOs

A total of 1,222 OR sequences were aligned and trimmed (see Section 2). The final length of the alignment was of 254 amino acid positions. To facilitate comparison, we retained the nomenclature used by Mitchell et al. (2019) to describe the phylogenetic groups and clades recovered in their phylogenetic analyses (i.e., groups 1, 2A, 2B, 3, 4, 5A, 5B, 6 and 7). Our results were overall congruent with those reported by Mitchell et al. (2019), with virtually all OR groups recovered with high support except group 6 and a different position for group 4, which was recovered as nested within group 3 (Figure 3a). While most of the genes fall into the same groups than in Mitchell et al. (2019), four genes (i.e., AplaOR1, PserOR120-121 and SlonOR34525c0g1) were not recovered for any of the previously proposed groups. The upregulated ORs in *S. longicornis* antennae were distributed among the different coleopteran OR groups, mostly clustered within group 1 and group 7 (with six and five genes, respectively). Exclusively expressed ORs in antennae are found in groups 1, 3, 4 and 7. The number of ORs is highly variable among these species (Figure 3b). *S. longicornis* and the other nonphytophagous species (i.e., *C. scutator*, *N. vespilloides*) exhibited relatively moderate OR repertoires and a similar distribution pattern (i.e., without representation in group 5A, moderate gene family expansions and the largest expansion in group 3; Figure 3b,d). All the ORCO sequences were recovered as a clade with high support and were used to root the tree, facilitating the identification of the ORCO candidate of *S. longicornis* (SlonORCO; Figure 3a,c).

The phylogeny of ORCOs (Figure 3c; with a final trimmed alignment of 478 amino acid positions) recovered clades for the different beetle families, but did not mirror the phylogeny of Coleoptera at the family level. For instance, all ORCOs of species of Cucujiformia were recovered in a clade and were subsequently clustered into their corresponding families. The same pattern was

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

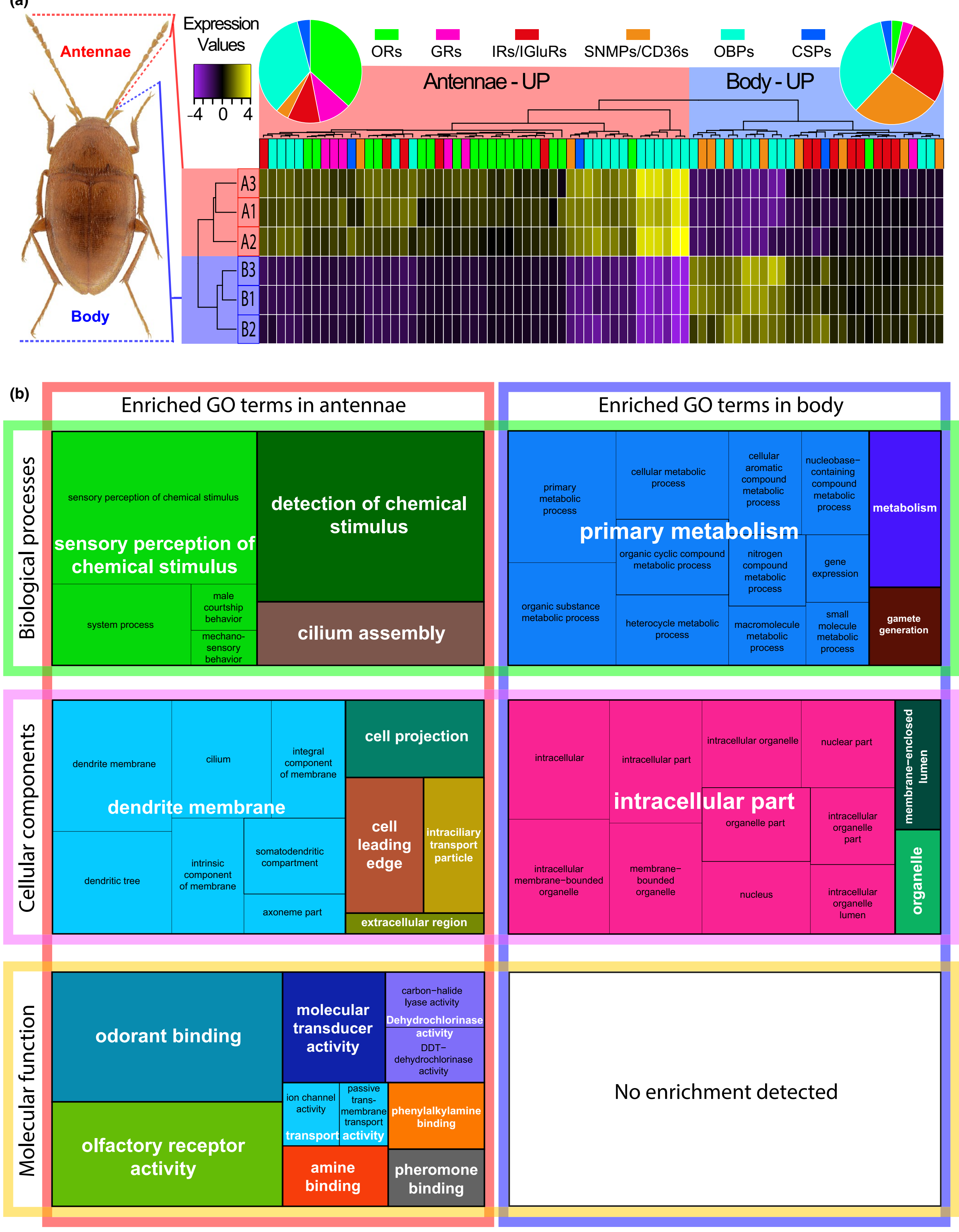
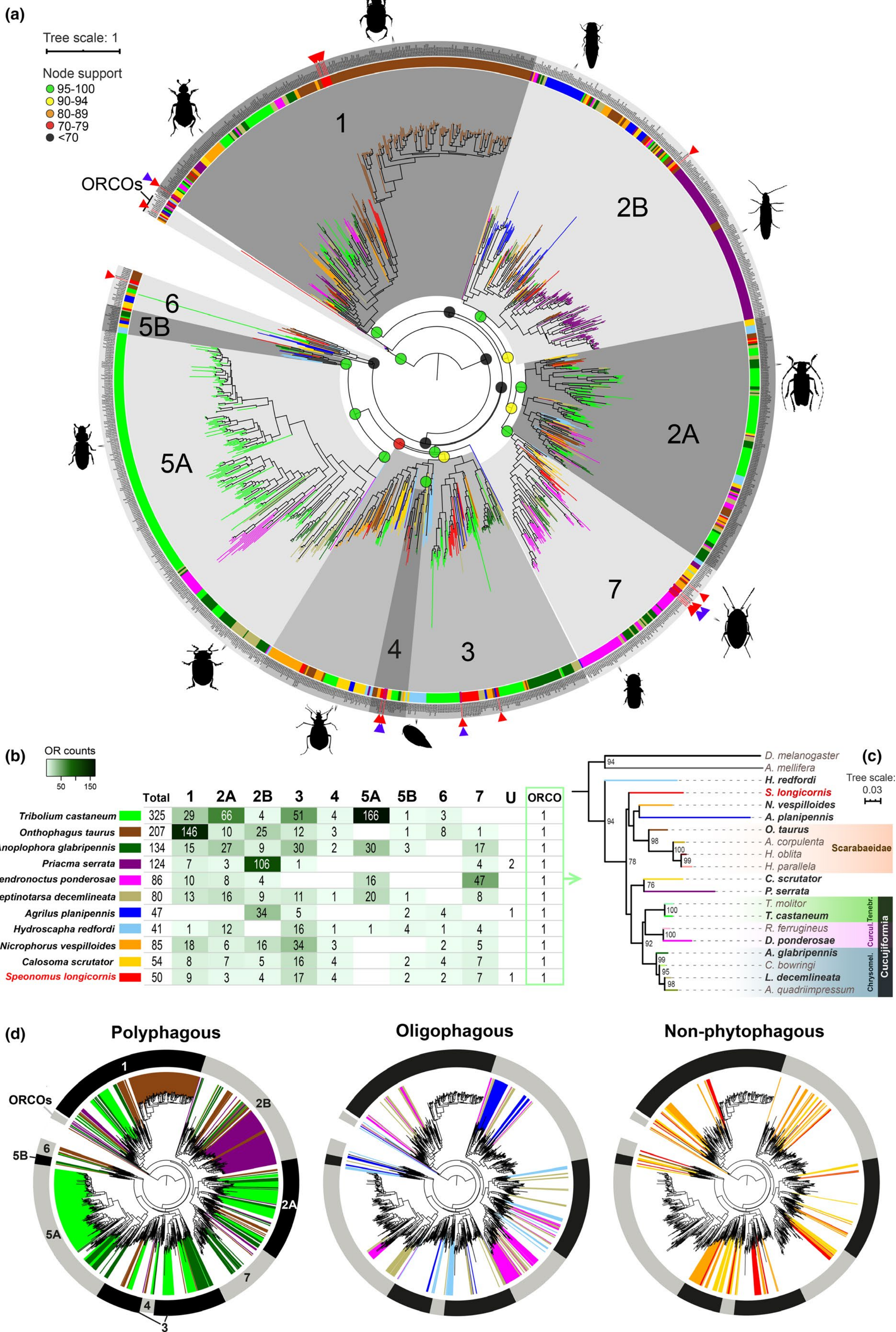


FIGURE 2 Differentially expressed genes in antennae and body in *S. longicornis*. (a) Heatmap of chemosensory genes of *S. longicornis* differentially expressed in antennae and the rest of the body. (b) Gene ontology (GO) treemaps for the differentially expressed genes in antennae versus the rest of the body. Biological process, molecular function and cellular component enriched GO terms are shown [Colour figure can be viewed at wileyonlinelibrary.com]

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



	Total	1	2A	2B	3	4	5A	5B	6	7	U	ORCO
Tribolium castaneum	325	29	66	4	51	4	166	1	3			1
Onthophagus taurus	207	146	10	25	12	3		1	8	1		1
Anoplophora glabripennis	134	15	27	9	30	2	30	3		17		1
Priacma serrata	124	7	3	106	1					4	2	1
Dendronoctus ponderosae	86	10	8	4			16			47		1
Leptinotarsa decemlineata	80	13	16	9	11	1	20	1		8		1
Agrilus planipennis	47			34	5			2	4		1	1
Hydroscapha redfordi	41	1	12		16	1	1	4	1	4		1
Nicrophorus vespilloides	85	18	6	16	34	3			2	5		1
Calosoma scrutator	54	8	7	5	16	4		2	4	7		1
Speonomus longicornis	50	9	3	4	17	4		2	2	7	1	1

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 3 Phylogeny of odourant receptors. (a) Maximum likelihood phylogenetic tree of odourant receptors (ORs) including OR sets of *S. longicornis* and other coleopterans from Mitchell et al. (2019), representing the proposed OR groups in grey ranges. Red triangles represent the upregulated genes in the antennae of *S. longicornis*. Purple triangles represent exclusively expressed genes in antennae. Species are colour coded as indicated in (b). (b) Number of OR genes of each OR group inferred for each species included in the phylogeny. "U" indicates unclassified ORs. (c) Maximum likelihood phylogeny of ORCO across coleopterans (see Methods and Table S1 for species codes). (d) Simplified representation of OR diversity recovered for each species, highlighting the OR repertoire of species with different feeding strategies [Colour figure can be viewed at wileyonlinelibrary.com]

observed for the ORCOs of the different species of Scarabaeidae. By contrast, the ORCO of *S. longicornis* did not cluster together with that of *N. vespilloides*, despite belonging to the same superfamily (i.e., Staphylinoidea).

3.5 | Phylogenetic inference of the annotated GRs

A total of 374 sequences were interrogated, resulting in a multiple sequence alignment of 271 amino acid positions. Notably, none of the candidate GRs from *S. longicornis* clustered together with GRs involved in the perception of fructose and other sugars in the other species (Figure 4). On the other hand, in several coleopteran GRs, including *S. longicornis*, we recovered three candidates that cluster with those of *D. melanogaster* involved in perception of CO_2, which are well characterized functionally (termed as GR1, GR2, GR3 in beetles, and GR21a, GR63a in *D. melanogaster*; Dippel et al., 2016; Jones et al., 2007; Kwon et al., 2007). One of these three candidate CO_2 receptors of *S. longicornis* was upregulated in antennae and was recovered as orthologous to the GR2 gene in beetles. We also identified a candidate bitter taste GR (SlonGR19567c0g1) that clustered together with strong support with previously identified as conserved bitter taste GRs for *A. planipennis* and *D. ponderosae* (Andersson et al., 2019). The rest of the genes were generally recovered in well-supported clades with species-specific differences in the extent of GR expansions. *S. longicornis* showed divergent GRs distributed along the tree exhibiting relatively small expansions (i.e., from two to five genes) and in general a relatively diminished gustatory repertoire, whereas the other species exhibit remarkable expansions and considerably larger repertoires. The functions of the other four upregulated GRs in antennae cannot be further predicted due to the lack of functional annotation of the genes they cluster together with.

3.6 | IR and IGluR phylogenies

Since IRs derive from IGluRs (see Introduction), our phylogenetic approach facilitated the initial annotations of both types of genes in *S. longicornis*. IGluRs are not directly associated with chemoreception but show a high sequence identity with the most conserved IRs (i.e., IR8a, IR25a). Three genes clustered together with N-methyl-D-aspartate receptors (NMDARs) and nine genes clustered with different IGluR clades, one upregulated in antennae (Figure 5a).

A total of 164 IR sequences were aligned and trimmed resulting in a multiple sequence alignment of 356 amino acid positions (Figure 5c). Several IRs of *S. longicornis* clustered together with conserved IRs in insects (i.e., IR8a, IR25a, IR93a, IR76b, IR21a, IR68a, IR40a, IR100, IR60a). No genes clustering together with IR41a were detected for *S. longicornis*. Several putative gene duplications were detected for *S. longicornis* (containing each from 2 to 5 IRs), more moderate in size than the large gene family expansions observed for *A. planipennis* and *D. ponderosae*, which included up to eight and 17 IRs, respectively. Our analyses revealed a gene duplication in IR25a, a highly conserved single copy gene virtually in all protostomes (with the exception of the parasitoid wasps *Nasonia vitripennis* and *Microplitis mediator* and the limpet *Lottia gigantea*, see Section 4). The copy of IR25a exhibiting the shortest branch in *S. longicornis* was upregulated in antennae (Figure 5a–c). In order to assess the robustness of our results, we followed four steps. First, all isoforms from both genes were visually inspected in an alignment (Figure S4), observing notable differences between the sequences belonging to the different genes and a high similarity for the isoforms of the same gene. Second, we used the expression values of each isoform in both genes to test whether only a low number of sequence reads from some of the isoforms were mapping back to one of the genes (which would indicate errors in the assembly), and tested if the expression levels were statistically different by means of a one-way ANOVA (Table S5). Expression levels were uniformly distributed across replicates with the exception of one isoform that was highly expressed in one replicate, indicating that the inference of the two genes was not a methodological artefact. Third, the final alignment for the phylogeny including all the taxa (using the longest isoforms as described in Section 2) was examined to test that they were not nonoverlapping fragmented genes. Fourth, both IR25a copies were annotated with HMMER to inspect the similarity of the domain profiles. These IR25a candidates (i.e., SlonIR11039c0g2 and SlonIR14393c1g1) shared 44% of identical residues whereas the conserved copy of *S. longicornis* (SlonIR14393c1g1) had between 69% to 72% of amino acid sequence identity with the IR25 candidates of the other coleopteran species. In addition, the protein annotation of the IR25a candidates of *S. longicornis* by HMMER resulted in highly similar domain profiles, suggesting their similarity at the structural level.

3.7 | OBP phylogeny

A total of 137 OBP sequences were included to explore the OBPs diversity of *S. longicornis*, resulting in a multiple sequence alignment of 110 amino acid positions after trimming. Our results suggest that OBPs in *S. longicornis* are relatively abundant compared to the other species, being the most diverse repertoire of this comparison after

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

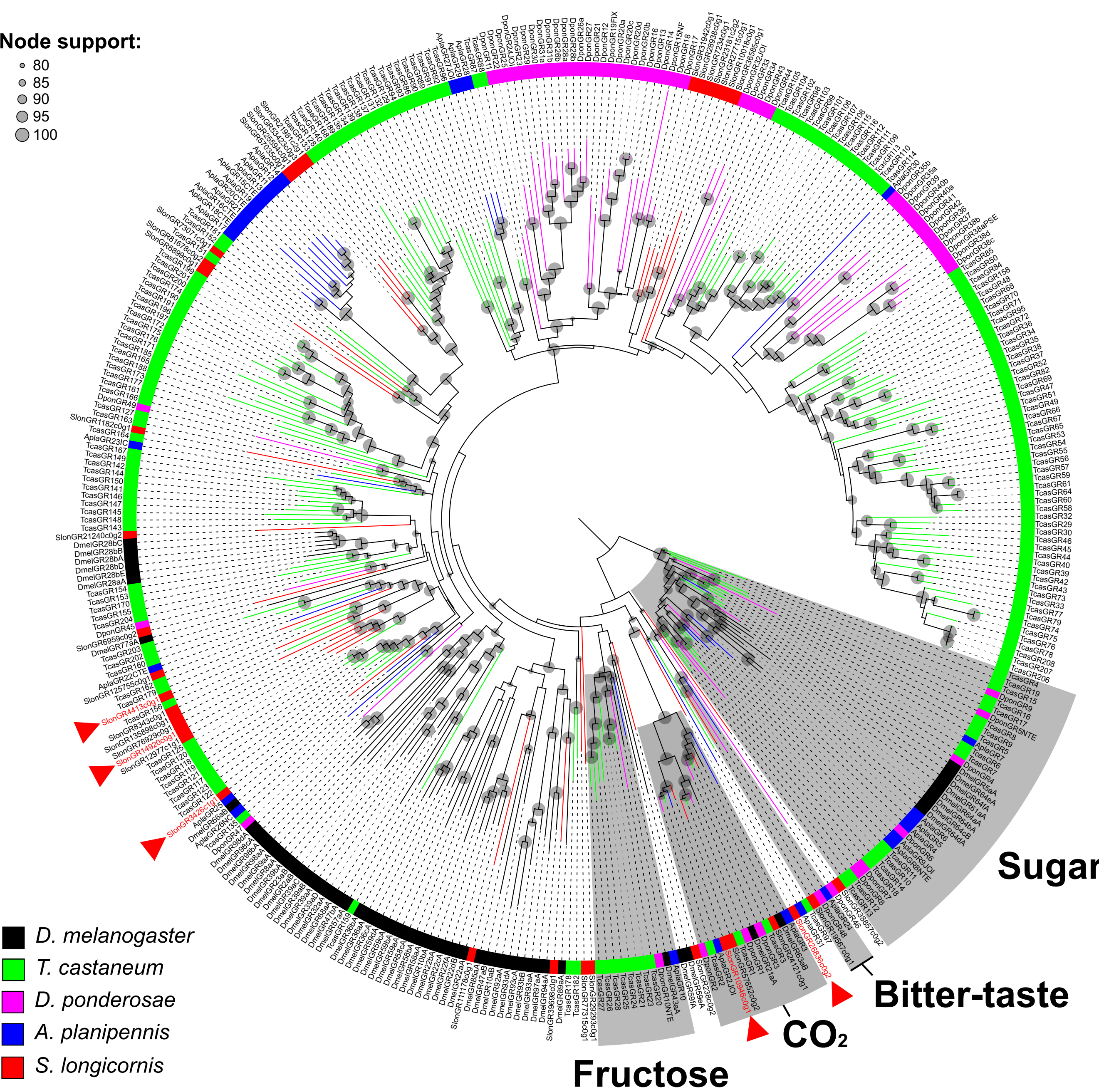
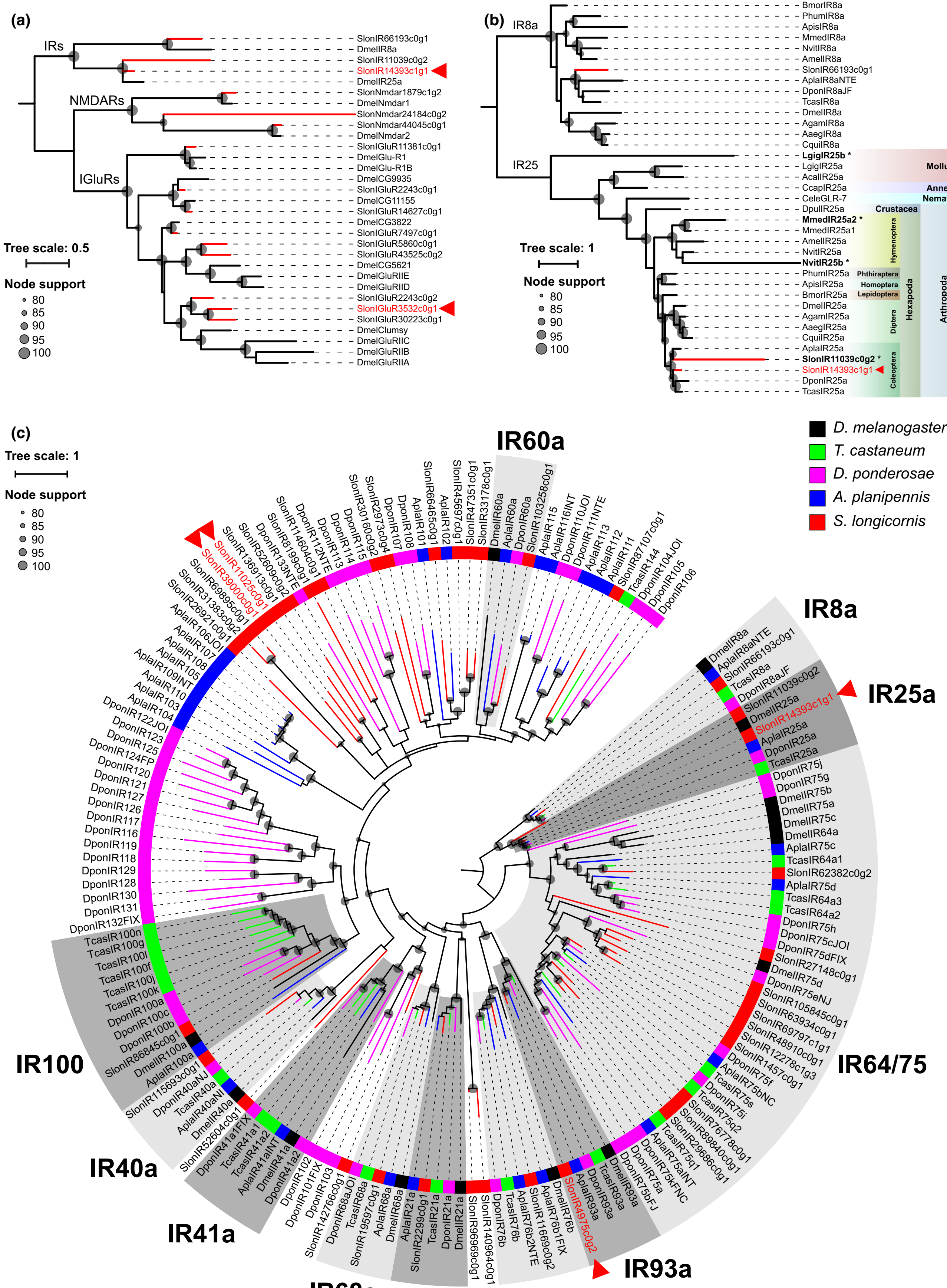


FIGURE 4 Phylogeny of gustatory receptors. Maximum likelihood phylogenetic tree of gustatory receptors (GRs) including GR sets of *S. longicornis*, other coleopterans from Andersson et al. (2019) and conserved GR sequences of *D. melanogaster*. Grey ranges represent well supported GR clades, indicating the proposed functions in the other species. Red triangles represent upregulated genes in the antennae of *S. longicornis* [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 5 Phylogeny of ionotropic receptors. (a) Maximum likelihood phylogenetic tree of ionotropic glutamate receptors (IGluRs) including sequences of *S. longicornis* and *D. melanogaster*. (b) Maximum likelihood phylogenetic tree of the ionotropic receptors clades IR25a and IR8a (the later used to root the tree), including the candidate genes for *S. longicornis* (*Slon*), beetle sequences retrieved from Andersson et al. (2019) and Dippel et al. (2016) and sequences from an expanded invertebrate taxon sampling obtained from Croset et al. (2010) and Wang et al. (2015). Species codes are as follows: *D. melanogaster* (*Dmel*), *Aedes aegypti* (*Aaeg*), *Culex quinquefasciatus* (*Cqui*), *Anopheles gambiae* (*Agam*), *Bombyx mori* (*Bmor*), *Apis mellifera* (*Amel*), *Nasonia vitripennis* (*Nvit*), *Microplitis mediator* (*Mmed*), *Acyrthosiphon pisum* (*Apis*), *Pediculus humanus* (*Phum*), *Daphnia pulex* (*Dpul*), *Caenorhabditis elegans* (*Cele*), *Capitella capitata* (*Ccap*), *Aplysia californica* (*Acal*) *and Lottia gigantea* (*Lgig*). Asterisks indicate the divergent copies of IR25 candidates. (c) Maximum likelihood phylogenetic tree of ionotropic receptors (IRs) including sequences of *S. longicornis*, other coleopterans from Andersson et al. (2019) and sequences of *D. melanogaster*. Grey ranges represent the conserved IR clades, based on the annotations of the other species. In all trees, red triangles represent upregulated genes in the antennae of *S. longicornis* [Colour figure can be viewed at wileyonlinelibrary.com]

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 6 Phylogeny of OBPs. Maximum likelihood phylogenetic tree of OBPs including sequences of *S. longicornis* and other coleopterans from Andersson et al. (2019). Grey ranges represent the main OBP clades described in previous studies. Red triangles represent upregulated genes in the antennae of *S. longicornis* [Colour figure can be viewed at wileyonlinelibrary.com]

T. castaneum (Figure 6). Several OBP candidates of *S. longicornis* clustered together with the OBP subgroups described for the other species in Andersson et al. (2019; i.e., classic-OBPs, minus-C, plus-C and antennal binding proteins II [ABPII]). Only seven out of the 17 OBPs upregulated in antennae correspond to the ABPII clade. Furthermore, two relatively large OBP expansions include the majority of the upregulated OBPs, formed by (i) the minus-C clade (with three upregulated genes in antennae), and (ii) a specific *S. longicornis* OBP lineage of ten genes (with five upregulated genes in the antennae).

The phylogenomic characterization of other gene families (SNMP/CD36 and CSPs) was also explored. The results and discussion are included as Supporting Information (Results S1).

4 | DISCUSSION

A highly complete transcriptome for the cave-dwelling beetle *Speonomus longicornis* was generated in the present study (Figure

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

S1). Combining the differential gene expression and GO enrichment analyses with a curated annotation pipeline for the chemosensory related genes, we were able to explore the chemosensory gene repertoire of *S. longicornis*. Furthermore, the phylogenetic inferences for each of the chemosensory gene families offered the opportunity to compare the repertoire of genes involved in chemosensation in *S. longicornis* to other beetle species that occupy a wide variety of ecological niches, all of them in surface habitats.

The differential gene expression (Figure 2a) and GO enrichment analysis (Figure 2b) allowed us to identify upregulated genes in antennae (where chemosensory structures—sensilla—are highly concentrated) and compare the overall enriched functions in the antennae versus the rest of the body. As expected, olfaction was recovered as the most prominent function in antennae, representing more than half of the enriched terms in the molecular function category (odourant-binding and olfactory reception activities), indicating that ORs and OBPs are playing a major role in how cave beetles receive and process airborne cues. ORCO was upregulated in antennae, as expected since it is an essential component of the functional heterodimers that facilitate odourant reception combined with other ORs (Stengl & Funk, 2013). In addition, only seven ORs were observed as exclusively expressed in antennae (Table S4), suggesting that this gene family may include genes with high specificity in these appendages. All in all, these results highlight the importance of antennae in odourant perception in this cave beetle; further gene expression studies including additional structures such as mouth appendages would give more detailed insights for the rest of the identified ORs.

Concerning the OR phylogeny (Figure 3), our results are mostly consistent with what was found in Mitchell et al. (2019). Large expansions in several OR groups are highlighted in the polyphagous/generalist species *T. castaneum*, *O. taurus*, *A. glabripennis* and *P. serrata*, whereas the oligophagous/specialist *D. ponderosae*, *L. decemlineata*, *A. planipennis* and *H. redfordi* exhibit relatively reduced OR repertoires. Therefore, an apparent correlation between the host breadth and the OR diversity of herbivore Coleoptera is observed, clearly exemplified by the extent and distribution of OR diversity in the wood boring species (i.e., *A. glabripennis*, *D. ponderosae*, *A. planipennis*; Andersson et al., 2019). The insectivorous *C. scrutator* and the scavengers *N. vespilloides* and *S. longicornis* could be perceived as polyphagous/generalists but they showed an apparent low number of ORs and relatively smaller expansions compared to the rest of the polyphagous herbivores. Our results suggest a relatively reduced OR repertoire of *S. longicornis* compared to the other species, which may result from adaptation to the deep subterranean environment conditions: very limited in primary production, oligotrophic and inhabited by a relatively smaller number of species compared to surface habitats. Moreover the air remains still, saturated with water vapour and the potential evaporation rate is negligible for long time periods (Howarth & Moldovan, 2018). These features suggest an homogeneous habitat, probably less diverse in airborne odourant clues than surface habitats. An extreme contraction of chemosensory gene repertoire, particularly for ORs and OBPs, was observed in the fig wasp *Ceratosolen solmsi* when compared to other hymenopterans, possibly reflecting its high host-specificity (Xiao et al., 2013). Further research comparing high quality transcriptomes and genomes of surface and subterranean species or including a high quality reference genome for the target species would help to validate this conclusion.

Regarding gustatory perception (Figure 4), five GRs were significantly enriched in the antennae in *S. longicornis*, indicating a substantial gustatory role in these appendages. This result is consistent with what was found for GR expression levels of *T. castaneum*, where similar values in the maxillary palps and the antennae were reported (Dippel et al., 2016). Remarkably, no GRs associated with the perception of fructose and other sugars were detected in *S. longicornis* (at least clustering together with functionally annotated genes in *D. melanogaster*) indicating that either *S. longicornis* does not have receptors for these types of carbohydrates, or their evolutionary origin is different from that in other beetles. Further research including transcriptomic data of other structures involved in taste perception (e.g., mouth appendages) would help to better sustain the absence of these highly conserved sugar receptors with the current approach. In insects, the entire loss of sugar receptors has only been documented in some obligate blood feeders (i.e., *Glossina morsitans*, *Cimex lectularius* and *Pediculus humanus*; Benoit et al., 2016; Kirkness et al., 2010; Obiero et al., 2014). Further comparative studies including nonphytophagous beetles inhabiting surface habitats would help to test the hypothesis that a lack of sugar receptors may be directly associated with a strict subterranean lifestyle.

CO_2 perception may be crucial for *S. longicornis* to orientate within its habitat and to detect decomposing organic matter in the darkness, the main food source for this species. We detected three candidate GRs clustering together with highly conserved CO_2 receptors of insects (Robertson & Kent, 2009), among which only one candidate was significantly expressed in antennae (Figure 4). Our results suggest that CO_2 perception may not be restricted to a single chemosensory structure, congruent with what was found in *T. castaneum* after comparing different body structures (Dippel et al., 2016). These results in beetles are in contrast to what was found in well studied dipterans. For instance, *D. melanogaster* has only two CO_2 receptors that form functional heteromers significantly enriched in antennae (i.e., DmelGR21a and DmelGR63a; Dippel et al., 2016; Jones et al., 2007; Kwon et al., 2007), whereas *A. gambiae* has three CO_2 receptors that are upregulated in the mouthparts (AgamGR22-24; Pitts et al., 2011). Further studies exploring differential gene expression in different body parts are needed to deepen our understanding of CO_2 perception in *S. longicornis*.

The GR repertoire of *S. longicornis* was small and similar in size to that observed in the oligophagous *A. planipennis* (ash tree specialist). Our findings on the gustatory perception of this cavernicolous beetle may reflect the poor diversity of gustatory substances in the hypogean habitat compared to surface environments. A notably diminished odourant and gustatory capabilities were reported in *Drosophila sechellia*, an endemic species to the Seychelles that is specialized in feeding on a single plant species (McBride, 2007). Similar tendencies were found in independently specialized drosophilids

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(McBride & Arguello, 2007) which showed a remarkable acceleration of gene loss when compared to other species. In contrast, the most remarkable GR family expansion in insects was found in *Periplaneta americana* (i.e. 522 GRs, with the most expanded clade (329 GRs) corresponding to potential bitter receptors), that been related to its omnivorous and opportunistic feeding habits. Moreover this species has also the most expanded OR repertoire of Blattodea species (Li et al., 2018). A large expansion of GRs has also been reported in the polyphagous moth *Helicoverpa armigera* compared to other lepidopterans, again finding bitter receptor GRs as the most expanded clade of the family (Xu et al., 2016).

Contrary to the specific role of IGluRs in synaptic communication, IRs have more diverse roles which in insects are often related to chemoreception (Koh et al., 2014; Rytz et al., 2013). While the most conserved IRs (e.g., IR8a and IR25a) act as coreceptors conferring multiple odour-evoked electrophysiological responses, more recently some insect IRs have been found to mediate specific stimulus forming heterodimers with more selectively expressed IR subunits (Abuin et al., 2011, 2019). For instance, in *D. melanogaster*, the highly conserved coreceptors IR93a and IR25a are coexpressed with IR21a, mediating physiological and behavioral responses to low temperatures (Knecht et al., 2017; Ni et al., 2016). In *S. longicornis* we found overexpression in the antennae of the candidate genes clustering together to *D. melanogaster* IR93a and IR25a, while the candidate IR21a (SlonIR2299c0 g1) was significantly underexpressed in these appendages (Figure 5c). By contrast, higher expression levels for IR21a of *D. melanogaster* were found in the antennae (Sánchez-Alcañiz et al., 2018) and the same was found for other Coleoptera (Bin et al., 2017; Dippel et al., 2016). However, we did not explore differential gene expression in different structures from the body, which could explain the differences in the observed results.

Remarkably, our results suggest a putative duplication in the two genes annotated as IR25a, despite being a highly conserved gene with a single copy in virtually all protostomes (Figure 5b). Duplications have only been reported for the limpet *L. gigantea* and for two parasitoid wasp species: *N. vitripennis* and *M. mediator* (Croset et al., 2010; Wang et al., 2015). The inferred phylogeny suggested different origins for the observed duplication in the IR25a candidates. A lineage-specific duplication has been observed in *N. vitripennis* and *M. mediator*, where both copies of IR25a were retrieved as sister to each other, suggesting either recent duplication or gene conversion. The candidate duplication in *S. longicornis* could represent an ancestral duplication in Coleoptera that was retained in this cave beetle or, alternatively, a recent gene duplication followed by extensive sequence divergence. Our results represent the first report of a gene duplication observed in this highly conserved gene in Coleoptera, which may indicate that the evolutionary history of IR25a and its role in chemoreception may be more complex than originally considered across arthropods.

Cave beetles inhabit a medium where air tends to be still and the ambient temperature and humidity fluctuate only by tiny amounts over long periods, and therefore a sensitive thermal detection may have a selective advantage. On the contrary, it may also result in a loss of selection of thermal detection ability and/or a loss of ability to adapt to thermal extremes (i.e., climate variability hypothesis; Stevens, 1989). Physiological experiments on a close relative species (i.e., the cave-dwelling *Speophyes lucidulus*, Leiodidae, Cholevinae) revealed an extreme sensitivity to small changes in temperature incurred by antennal receptors (Corbière-Tichané & Loftus, 1983) that may be mediated by some of the inferred candidate IRs. Consequently, other relevant IRs for *S. longicornis* may be those potentially related to humidity sensing. The functionally characterized IR40a and IR68a in *D. melanogaster* have been seen to be coexpressed with IR93a and IR25a in specialized sensory neurons of the antennae performing hygrosensory responses (Enjin, 2017; Knecht et al., 2017). Through the phylogenetic analysis we identified the hygroreceptor candidates (IR40a and IR68a) for *S. longicornis* (Figure 5c), although we did not find significant differences in the expression values between antennae and the rest of the body. The rest of the candidate IRs annotated in *S. longicornis* (i.e., IR8a, IR76b, IR75a, IR64a, 100a and IR60a; see Figure 5c; Table S2) have been shown to be potentially involved in taste and odour transduction in *D. melanogaster*, suggesting candidate odourant and gustatory roles in *S. longicornis*.

Regarding OBPs, the GO enrichment analysis highlighted odourant-binding functions vastly enriched in antennae (Figure 2b). In addition, the differential gene expression analyses identified a high number of upregulated genes (17 in antennae; Figure 2a). The large number of OBPs annotated in *S. longicornis* suggests a relatively diverse repertoire with species-specific gene duplications and gene family expansions; this may indicate an important role of these proteins in odourant perception in this subterranean beetle. Notably, less than half of the upregulated OBPs in antennae clustered together with the previously described "antennal binding proteins II" (ABPII) in Vieira and Rozas (2011; Figure 6). These results indicate that although ABPII were described as OBPs typically enriched in antennae, some genes of this clade may also be differentially expressed in some other body structures, as also found for *T. castaneum* (Dippel et al., 2014). The rest of the upregulated OBPs in both conditions clustered together with the different OBP groups described in previous studies (Andersson et al., 2019; Dippel et al., 2014; Vieira & Rozas, 2011). Further research including closer relatives to this cave species with different ecological preferences will allow us to test the hypothesis that subterranean specialization has modified the chemosensory capabilities in Coleoptera.

5 | CONCLUSION

In this study, we characterized for the first time the chemosensory gene repertoire of an obligate subterranean species, the cave-dwelling coleopteran *S. longicornis*. We found relatively diminished odourant and gustatory repertoires compared to polyphagous coleopterans inhabiting surface habitats and more similar to species considered specialists based on their feeding habits. Considering the selective pressures of the niche that *S. longicornis* occupies (i.e.,

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

limited resources, poor diversity and heterogeneous distribution of food, among others), an optimized chemosensory repertoire in terms of diversity may result from its adaptation to the deep subterranean environment. In this obligate cave-dwelling beetle, we identified some putative gene losses (e.g., sugar GRs and IR41a). Furthermore, several gene duplications and gene family expansions were observed. A putative duplication of the gene IR25a was identified, which might potentially have facilitated adaptation to subterranean conditions in this cave beetle. Our study thus paves the way towards a better understanding of how subterranean animals perceive their particular environment.

ACKNOWLEDGEMENTS

This work was supported by the Ministerio de Economía y Competitividad and the Ministerio de Ciencia of Spain (CGL2016-76705-P to Ignacio Ribera, PID2019-108824GA-I00 to Rosa Fernández, and CGL2016-75255 and PID2019-103947GB to Julio Rozas). Pau Balart-García and Paula Escuer were supported by an FPI grant (Ministerio de Economía y Competitividad BES-2017-081050 and BES-2017-081740, respectively). Rosa Fernández was supported by a Marie Sklodowska-Curie grant (747607) and a Ramón y Cajal fellowship (Ministerio de Economía y Competitividad, RyC2017-22492). C. Vanderbergh and E. Ruzzier kindly provided the images for Figure 1a,b, respectively; the latter was obtained at the Electron Microscopy Service of the Institute of Marine Sciences (ICM) of Barcelona, Spain. This article is dedicated to the memory of our wonderful colleague, Ignacio Ribera, who worked intensively on this project until his very last days and recently sadly passed away.

AUTHOR CONTRIBUTIONS

Pau Balart-García conceived the study, generated, interpreted and analysed the data, prepared the figures and tables and wrote the first version of the manuscript. Alexandra Cieslak conceived the study, processed the samples and generated the data. Paula Escuer assisted in data analysis and interpretation. Julio Rozas provided computational resources and assisted in data interpretation. Ignacio Ribera conceived the study, provided resources and interpreted the data. Rosa Fernández conceived and supervised the study, analysed and interpreted the data, and wrote the first version of the manuscript. All authors contributed to the final version of the manuscript.

DATA AVAILABILITY STATEMENT

Raw reads have been deposited in the National Center for Biotechnology Information (NCBI; BioProject accession number PRJNA667243, BioSample accession number SAMN16362632). Assembled sequences, alignments and custom scripts have been deposited in github (https://github.com/MetazoaPhylogenomicsLab/Balart_Garcia_et_al_2021_MolecularEcology).

ORCID

Pau Balart-García https://orcid.org/0000-0002-8292-4674
Alexandra Cieslak https://orcid.org/0000-0003-3002-7995
Paula Escuer https://orcid.org/0000-0002-5941-0106
Julio Rozas https://orcid.org/0000-0002-6839-9148
Ignacio Ribera https://orcid.org/0000-0002-2791-7615
Rosa Fernández https://orcid.org/0000-0002-4719-6640

REFERENCES

Abuin, L., Bargeton, B., Ulbrich, M. H., Isacoff, E. Y., Kellenberger, S., & Benton, R. (2011). Functional architecture of olfactory ionotropic glutamate receptors. *Neuron*, *69*(1), 44–60.

Abuin, L., Prieto-Godino, L. L., Pan, H., Gutierrez, C., Huang, L., Jin, R., & Benton, R. (2019). In vivo assembly and trafficking of olfactory Ionotropic Receptors. *BMC Biology*, *17*(1), 34.

Almudi, I., Vizueta, J., Wyatt, C. D. R., de Mendoza, A., Marlétaz, F., Firbas, P. N., Feuda, R., Masiero, G., Medina, P., Alcaina-Caro, A., Cruz, F., Gómez-Garrido, J., Gut, M., Alioto, T. S., Vargas-Chavez, C., Davie, K., Misof, B., González, J., Aerts, S., … Casares, F. (2020). Genomic adaptations to aquatic and aerial life in mayflies and the origin of insect wings. *Nature Communications*, *11*(1), 2631.

Al-Shahrour, F., Minguez, P., Tárraga, J., Medina, I., Alloza, E., Montaner, D., & Dopazo, J. (2007). FatiGO +: A functional profiling tool for genomic data. Integration of functional annotation, regulatory motifs and interaction data with microarray experiments. *Nucleic Acids Research*, *35*(Web Server issue), W91–W96.

Andersson, M. N., Keeling, C. I., & 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*(1), 690.

Anholt, R. R. H. (2020). Chemosensation and evolution of Drosophila host plant selection. *iScience*, *23*(1), 100799.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, *57*(1), 289–300.

Benoit, J. B., Adelman, Z. N., Reinhardt, K., Dolan, A., Poelchau, M., Jennings, E. C., Szuter, E. M., Hagan, R. W., Gujar, H., Shukla, J. N., Zhu, F., Mohan, M., Nelson, D. R., Rosendale, A. J., Derst, C., Resnik, V., Wernig, S., Menegazzi, P., Wegener, C., … Richards, S. (2016). Unique features of a global human ectoparasite identified through sequencing of the bed bug genome. *Nature Communications*, *7*(1), 10165.

Benton, R. (2015). Multigene family evolution: Perspectives from insect chemoreceptors. *Trends in Ecology & Evolution*, *30*(10), 590–600.

Benton, R., Vannice, K. S., Gomez-Diaz, C., & Vosshall, L. B. (2009). Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell*, *136*(1), 149–162.

Bin, S.-Y., Qu, M.-Q., Li, K.-M., Peng, Z.-Q., Wu, Z.-Z., & Lin, J.-T. (2017). Antennal and abdominal transcriptomes reveal chemosensory gene families in the coconut hispine beetle, *Brontispa longissima*. *Scientific Reports*, *7*(1), 2809.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120.

Brand, P., Robertson, H. M., Lin, W., Pothula, R., Klingeman, W. E., Jurat-Fuentes, J. L., & Johnson, B. R. (2018). The origin of the odorant receptor gene family in insects. *eLife*, *7*, e38340.

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.

Cartwright, R. A., Schwartz, R. S., Merry, A. L., & Howell, M. M. (2017). The importance of selection in the evolution of blindness in cavefish. *BMC Evolutionary Biology*, *17*(1), 45.

Cieslak, A., Fresneda, J., & Ribera, I. (2014). Life-history specialization was not an evolutionary dead-end in Pyrenean cave beetles. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1781), 20132978.

Clyne, P. J., Warr, C. G., Freeman, M. R., Lessing, D., Kim, J., & Carlson, J. R. (1999). A novel family of divergent seven-transmembrane

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

proteins: Candidate odorant receptors in *Drosophila*. *Neuron*, *22*(2), 327–338.

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., & de Hoon, M. J. L. (2009). Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, *25*(11), 1422–1423.

Corbière-Tichané, G., & Loftus, R. (1983). Antennal thermal receptors of the cave beetle *Speophyes lucidulus* Delar. *Journal of Comparative Physiology*, *153*(3), 343–351.

Croset, V., Rytz, R., Cummins, S. F., Budd, A., Brawand, D., Kaessmann, H., Gibson, T. J., & Benton, R. (2010). Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genetics*, *6*(8), e1001064.

Culver, D. C., & Pipan, T. (2019). *The biology of caves and other subterranean habitats*. Oxford University Press.

Delay, B. (1978). Milieu souterrain et ecophysiologie de la reproduction et du développement des coléoptères Bathysciinae hypoges. *Mémoires De Biospéologie*, *5*, 1–349.

Deleurance, S. (1963). Recherches sur les coléoptères troglobies de la sous-famille des Bathysciinae. *Annales Des Sciences Naturelles (Paris) Zoologie*, *1*, 1–172.

Dippel, S., Kollmann, M., Oberhofer, G., Montino, A., Knoll, C., Krala, M., Rexer, K.-H., Frank, S., Kumpf, R., Schachtner, J., & Wimmer, E. A. (2016). Morphological and transcriptomic analysis of a beetle chemosensory system reveals a Gnathal Olfactory Center. *BMC Biology*, *14*(1), 90.

Dippel, S., Oberhofer, G., Kahnt, J., Gerischer, L., Opitz, L., Schachtner, J., Stanke, M., Schütz, S., Wimmer, E. A., & Angeli, S. (2014). Tissue-specific transcriptomics, chromosomal localization, and phylogeny of chemosensory and odorant binding proteins from the red flour beetle *Tribolium castaneum* reveal subgroup specificities for olfaction or more general functions. *BMC Genomics*, *15*(1), 1141.

Enjin, A. (2017). Humidity sensing in insects-from ecology to neural processing. *Current Opinion in Insect Science*, *24*, 1–6.

Eyun, S.-I., Soh, H. Y., Posavi, M., Munro, J. B., Hughes, D. S. T., Murali, S. C., Qu, J., Dugan, S., Lee, S. L., Chao, H., Dinh, H., Han, Y. I., Doddapaneni, H. V., Worley, K. C., Muzny, D. M., Park, E.-O., Silva, J. C., Gibbs, R. A., Richards, S., & Lee, C. E. (2017). Evolutionary history of chemosensory-related gene families across the Arthropoda. *Molecular Biology and Evolution*, *34*(8), 1838–1862.

Fernández, R., & Gabaldón, T. (2020). Gene gain and loss across the metazoan tree of life. *Nature Ecology & Evolution*, *4*(4), 524–533.

Gao, Q., & Chess, A. (1999). Identification of candidate *Drosophila* olfactory receptors from genomic DNA sequence. *Genomics*, *60*(1), 31–39.

Glaçon, S. (1953). The evolutive cycle of a cave coleopter *Speonomus longicornis* Saulcy. *Comptes Rendus Hebdomadaires Des Seances De L'academie Des Sciences*, *236*(25), 2443–2445.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652.

Grimaldi, D., & Engel, M. S. (2005). *Evolution of the insects*. Cambridge University Press.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B. O., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*(2), 518–522.

Howarth, F. G., & Moldovan, O. T. (2018). Where cave animals live. In O. T. Moldovan, Ľ. Kováč, & S. Halse (Eds.), *Cave ecology* (pp. 23–37). Springer International Publishing.

Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., von Mering, C., & Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Molecular Biology and Evolution*, *34*(8), 2115–2122.

Jeannel, R. (1924). Monographie des Bathysciinae. *Archives De Zoologie Expérimentale Et Générale (Paris)*, *63*, 1–436.

Jones, W. D., Cayirlioglu, P., Kadow, I. G., & Vosshall, L. B. (2007). Two chemosensory receptors together mediate carbon dioxide detection in *Drosophila*. *Nature*, *445*(7123), 86–90.

Joseph, R. M., & Carlson, J. R. (2015). *Drosophila* chemoreceptors: A molecular interface between the chemical world and the brain. *Trends in Genetics*, *31*(12), 683–695.

Kirkness, E. F., Haas, B. J., Sun, W., Braig, H. R., Perotti, M. A., Clark, J. M., Lee, S. H., Robertson, H. M., Kennedy, R. C., Elhaik, E., Gerlach, D., Kriventseva, E. V., Elsik, C. G., Graur, D., Hill, C. A., Veenstra, J. A., Walenz, B., Tubio, J. M. C., Ribeiro, J. M. C., … Pittendrigh, B. R. (2010). Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(27), 12168–12173.

Knecht, Z. A., Silbering, A. F., Cruz, J., Yang, L., Croset, V., Benton, R., & Garrity, P. A. (2017). Ionotropic receptor-dependent moist and dry cells control hygrosensation in Drosophila. *eLife*, *6*(e26654), 1–11. https://doi.org/10.7554/elife.26654

Koh, T.-W., He, Z., Gorur-Shandilya, S., Menuz, K., Larter, N. K., Stewart, S., & Carlson, J. R. (2014). The *Drosophila* IR20a clade of ionotropic receptors are candidate taste and pheromone receptors. *Neuron*, *83*(4), 850–865.

Kwon, J. Y., Dahanukar, A., Weiss, L. A., & Carlson, J. R. (2007). The molecular basis of CO2 reception in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(9), 3574–3578.

Laetsch, D. R., & Blaxter, M. L. (2017). BlobTools: Interrogation of genome assemblies. *F1000Research*, *6*, 1287.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359.

Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Research*, *47*(W1), W256–W259.

Li, S., Zhu, S., Jia, Q., Yuan, D., Ren, C., Li, K., Liu, S., Cui, Y., Zhao, H., Cao, Y., Fang, G., Li, D., Zhao, X., Zhang, J., Yue, Q., Fan, Y., Yu, X., Feng, Q., & Zhan, S. (2018). The genomic and functional landscapes of developmental plasticity in the American cockroach. *Nature Communications*, *9*(1), 1008.

Luo, X.-Z., Antunes-Carvalho, C., Wipfler, B., Ribera, I., & Beutel, R. G. (2019). The cephalic morphology of the troglobiontic cholevine species *Troglocharinus ferreri* (Coleoptera, Leiodidae). *Journal of Morphology*, *280*(8), 1207–1221.

McBride, C. S. (2007). Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(12), 4996–5001.

McBride, C. S., & Arguello, J. R. (2007). Five *Drosophila* genomes reveal nonneutral evolution and the signature of host specialization in the chemoreceptor superfamily. *Genetics*, *177*, 1395–1416.

McKenna, D. D., Shin, S., Ahrens, D., Balke, M., Beza-Beza, C., Clarke, D. J., Donath, A., Escalona, H. E., Friedrich, F., Letsch, H., Liu, S., Maddison, D., Mayer, C., Misof, B., Murin, P. J., Niehuis, O., Peters, R. S., Podsiadlowski, L., Pohl, H., … Beutel, R. G. (2019). The evolution and

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

genomic basis of beetle diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(49), 24729–24737.

Mirarab, S., Nguyen, N., Guo, S., Wang, L.-S., Kim, J., & Warnow, T. (2015). PASTA: Ultra-large multiple sequence alignment for nucleotide and amino-acid sequences. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, *22*(5), 377–386.

Missbach, C., Dweck, H. K. M., Vogel, H., Vilcinskas, A., Stensmyr, M. C., Hansson, B. S., & Grosse-Wilde, E. (2014). Evolution of insect olfactory receptors. *eLife*, *3*(e02115), 1–22. https://doi.org/10.7554/elife.02115

Missbach, C., Vogel, H., Hansson, B. S., & Groβe-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. *Chemical Senses*, *40*(9), 615–626.

Mitchell, R. F., Schneider, T. M., Schwartz, A. M., Andersson, M. N., & McKenna, D. D. (2019). The diversity and evolution of odorant receptors in beetles (Coleoptera). *Insect Molecular Biology*, *29*(1), 77–91.

Nei, M., Niimura, Y., & Nozawa, M. (2008). The evolution of animal chemosensory receptor gene repertoires: Roles of chance and necessity. *Nature Reviews Genetics*, *9*(12), 951–963.

Nei, M., & Rooney, A. P. (2005). Concerted and birth-and-death evolution of multigene families. *Annual Review of Genetics*, *39*(1), 121–152.

Ngoc, P. C. T., Greenhalgh, R., Dermauw, W., Rombauts, S., Bajda, S., Zhurov, V., Grbić, M., Van de Peer, Y., Van Leeuwen, T., Rouzé, P., & Clark, R. M. (2016). Complex evolutionary dynamics of massively expanded chemosensory receptor families in an extreme generalist chelicerate herbivore. *Genome Biology and Evolution*, *8*(11), 3323–3339.

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*(1), 268–274.

Ni, L., Klein, M., Svec, K. V., Budelli, G., Chang, E. C., Ferrer, A. J., Benton, R., Samuel, A. D. T., & Garrity, P. A. (2016). The ionotropic receptors IR21a and IR25a mediate cool sensing in Drosophila. *eLife*, *5*(e13254), 1–12. https://doi.org/10.7554/eLife.13254

Nichols, Z., & Vogt, R. G. (2008). The SNMP/CD36 gene family in Diptera, Hymenoptera and Coleoptera: *Drosophila melanogaster*, *D. pseudoobscura*, *Anopheles gambiae*, *Aedes aegypti*, *Apis mellifera*, and *Tribolium castaneum*. *Insect Biochemistry and Molecular Biology*, *38*(4), 398–415.

Obiero, G. F. O., Mireji, P. O., Nyanjom, S. R. G., Christoffels, A., Robertson, H. M., & Masiga, D. K. (2014). Odorant and gustatory receptors in the tsetse fly *Glossina morsitans morsitans*. *PLoS Neglected Tropical Diseases*, *8*(4), e2663.

Pallarés, S., Colado, R., Botella-Cruz, M., Montes, A., Balart-García, P., Bilton, D. T., Sánchez-Fernández, D. (2020). Loss of heat acclimation capacity could leave subterranean specialists highly sensitive to climate change. *Animal Conservation*. https://doi.org/10.1111/acv.12654. [Epub ahead of print].

Parzefall, J. (2001). A review of morphological and behavioural changes in the cave molly, *Poecilia mexicana*, from Tabasco, Mexico. *Environmental Biology of Fishes*, *62*, 263–275.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*(4), 417–419.

Pelosi, P., Iovinella, I., Felicioli, A., & Dani, F. R. (2014). Soluble proteins of chemical communication: An overview across arthropods. *Frontiers in Physiology*, *5*, 320.

Pelosi, P., Iovinella, I., Zhu, J., Wang, G., & Dani, F. R. (2018). Beyond chemoreception: Diverse tasks of soluble olfactory proteins in insects. *Biological Reviews of the Cambridge Philosophical Society*, *93*(1), 184–200.

Picelli, S., Björklund, Å. K., Faridani, O. R., Sagasser, S., Winberg, G., & Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nature Methods*, *10*(11), 1096–1098.

Pitts, R. J., Rinker, D. C., Jones, P. L., Rokas, A., & Zwiebel, L. J. (2011). Transcriptome profiling of chemosensory appendages in the malaria vector *Anopheles gambiae* reveals tissue- and sex-specific signatures of odor coding. *BMC Genomics*, *12*, 271.

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS One*, *5*(3), e9490.

Ribera, I., Fresneda, J., Bucur, R., Izquierdo, A., Vogler, A. P., Salgado, J. M., & Cieslak, A. (2010). Ancient origin of a Western Mediterranean radiation of subterranean beetles. *BMC Evolutionary Biology*, *10*, 29.

Rizzo, V., Sánchez-Fernández, D., Fresneda, J., Cieslak, A., & Ribera, I. (2015). Lack of evolutionary adjustment to ambient temperature in highly specialized cave beetles. *BMC Evolutionary Biology*, *15*, 10.

Robertson, H. M. (2019). Molecular evolution of the major arthropod chemoreceptor gene families. *Annual Review of Entomology*, *64*(1), 227–242.

Robertson, H. M., & Kent, L. B. (2009). Evolution of the gene lineage encoding the carbon dioxide receptor in insects. *Journal of Insect Science*, *9*(19), 1–14.

Robertson, H. M., Warr, C. G., & Carlson, J. R. (2003). Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(Suppl. 2), 14537–14542.

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140.

Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, *11*(3), R25.

Roys, C. (1954). Olfactory nerve potentials a direct measure of chemoreception in insects. *Annals of the New York Academy of Sciences*, *58*(2), 250–255.

Rytz, R., Croset, V., & Benton, R. (2013). Ionotropic receptors (IRs): Chemosensory ionotropic glutamate receptors in *Drosophila* and beyond. *Insect Biochemistry and Molecular Biology*, *43*(9), 888–897.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning: A laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sánchez-Alcañiz, J. A., Silbering, A. F., Croset, V., Zappia, G., Sivasubramaniam, A. K., Abuin, L., Sahai, S. Y., Münch, D., Steck, K., Auer, T. O., Cruchet, S., Neagu-Maier, G. L., Sprecher, S. G., Ribeiro, C., Yapici, N., & Benton, R. (2018). An expression atlas of variant ionotropic glutamate receptors identifies a molecular basis of carbonation sensing. *Nature Communications*, *9*(1), 4252.

Sánchez-Gracia, A., Vieira, F. G., & Rozas, J. (2009). Molecular evolution of the major chemosensory gene families in insects. *Heredity*, *103*(3), 208–216.

Schoonhoven, L. M., van Loon, J. J. A., & Dicke, M. (2005). *Insect-plant biology* (2nd ed., pp. 6–24). Oxford University Press.

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*(19), 3210–3212.

Stengl, M., & Funk, N. W. (2013). The role of the coreceptor Orco in insect olfactory transduction. *Journal of Comparative Physiology A*, *199*(11), 897–909.

Stevens, G. C. (1989). The latitudinal gradient in geographical range: How so many species coexist in the tropics. *The American Naturalist*, *133*(2), 240–256.

Stocker, R. F. (1994). The organization of the chemosensory system in *Drosophila melanogaster*: A rewiew. *Cell and Tissue Research*, *275*(1), 3–26.

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Stürckow, B. (1970). Responses of olfactory and gustatory receptor cells in insects. *Advances in Chemoreception*, *1*, 107–159.

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*, *6*(7), e21800.

Thoma, M., Missbach, C., Jordan, M. D., Grosse-Wilde, E., Newcomb, R. D., & Hansson, B. S. (2019). Transcriptome surveys in silverfish suggest a multistep origin of the insect odorant receptor gene family. *Frontiers in Ecology and Evolution*, *7*, 281.

Turk, S., Sket, B., & Sarbu, Ş. (1996). Comparison between some epigean and hypogean populations of *Asellus aquaticus* (Crustacea: Isopoda: Asellidae). *Hydrobiologia*, *337*(1–3), 161–170.

Vieira, F. G., & Rozas, J. (2011). Comparative genomics of the odorant-binding and chemosensory protein gene families across the arthropoda: Origin and evolutionary history of the chemosensory system. *Genome Biology and Evolution*, *3*, 476–490.

Vieira, F. G., Sánchez-Gracia, A., & Rozas, J. (2007). Comparative genomic analysis of the odorant-binding protein family in 12 *Drosophila* genomes: Purifying selection and birth-and-death evolution. *Genome Biology*, *8*(11), R235.

Vizueta, J., Escuer, P., Frías-López, C., Guirao-Rico, S., Hering, L., Mayer, G., Rozas, J., & Sánchez-Gracia, A. (2020). Evolutionary history of major chemosensory gene families across Panarthropoda. *Molecular Biology and Evolution*, *37*(12), 3601–3615.

Vizueta, J., Frías-López, C., Macías-Hernández, N., Arnedo, M. A., Sánchez-Gracia, A., & Rozas, J. (2016). Evolution of chemosensory gene families in arthropods: Insight from the first inclusive comparative transcriptome analysis across spider appendages. *Genome Biology and Evolution*, *9*(1), 178–196.

Vizueta, J., Sánchez-Gracia, A., & Rozas, J. (2020). BITACORA: A comprehensive tool for the identification and annotation of gene families in genome assemblies. *Molecular Ecology Resources*, *20*, 1445–1452.

Vogel, H., Shukla, S. P., Engl, T., Weiss, B., Fischer, R., Steiger, S., Heckel, D. G., Kaltenpoth, M., & Vilcinskas, A. (2017). The digestive and defensive basis of carcass utilization by the burying beetle and its microbiota. *Nature Communications*, *8*, 15186.

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A., & Axel, R. (1999). A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell*, *96*(5), 725–736.

Vosshall, L. B., & Stocker, R. F. (2007). Molecular architecture of smell and taste in *Drosophila*. *Annual Review of Neuroscience*, *30*, 505–533.

Wang, D., Pentzold, S., Kunert, M., Groth, M., Brandt, W., Pasteels, J. M., Boland, W., & Burse, A. (2018). A subset of chemosensory genes differs between two populations of a specialized leaf beetle after host plant shift. *Ecology and Evolution*, *8*(16), 8055–8075.

Wang, H.-C., Minh, B. Q., Susko, E., & Roger, A. J. (2018). Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Systematic Biology*, *67*, 216–235.

Wang, S.-N., Peng, Y., Lu, Z.-Y., Dhiloo, K. H., Gu, S.-H., Li, R.-J., & Guo, Y.-Y. (2015). Identification and expression analysis of putative chemosensory receptor genes in *Microplitis mediator* by antennal transcriptome screening. *International Journal of Biological Sciences*, *11*(7), 737–751.

Wu, Z., Bin, S., He, H., Wang, Z., Li, M., & Lin, J. (2016). Differential expression analysis of chemoreception genes in the striped Flea Beetle *Phyllotreta striolata* using a transcriptomic approach. *PLoS One*, *11*(4), e0153067.

Xu, W., Papanicolaou, A., Zhang, H.-J., & Anderson, A. (2016). Expansion of a bitter taste receptor family in a polyphagous insect herbivore. *Scientific Reports*, *6*, 23666.

Yamamoto, Y., Byerly, M. S., Jackman, W. R., & Jeffery, W. R. (2009). Pleiotropic functions of embryonic sonic hedgehog expression link jaw and taste bud amplification with eye loss during cavefish evolution. *Developmental Biology*, *330*(1), 200–211.

Yang, J., Chen, X., Bai, J., Fang, D., Qiu, Y., Jiang, W., Yuan, H., Bian, C., Lu, J., He, S., Pan, X., Zhang, Y., Wang, X., You, X., Wang, Y., Sun, Y., Mao, D., Liu, Y., Fan, G., … Shi, Q. (2016). The *Sinocyclocheilus* cavefish genome provides insights into cave adaptation. *BMC Biology*, *14*, 1.

Xiao, J. H., Yue, Z., Jia, L. Y., Yang, X. H., Niu, L. H., Wang, Z., Zhang, P., Sun, B. F., He, S. M., Li, Z., Xiong, T. L., Xin, W., Gu, H. F., Wang, B, Werren, J. H., Murphy, R. W, Wheeler, D., Niu, L. M., Ma, G. C., … Huang, D. W. (2013). Obligate mutualism within a host drives the extreme specialization of a fig wasp genome. *Genome Biology*, *14*(12), R141. http://dx.doi.org/10.1186/gb-2013-14-12-r141

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Balart-García P, Cieslak A, Escuer P, Rozas J, Ribera I, Fernández R. Smelling in the dark: Phylogenomic insights into the chemosensory system of a subterranean beetle. *Mol Ecol*. 2021;30:2573–2590. https://doi.org/10.1111/mec.15921

1365294x, 2021, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.15921 by Readcube (Labtiva Inc.), Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License