Highly dynamic evolution of the chemosensory gene repertoire driven by gene gain and expansion across subterranean beetles

Pau Balart-García[1*], Tessa M. Bradford[2,3], Perry G. Beasley-Hall[2,3], Slavko Polak[4], Ignacio Ribera[1], Steven J. B. Cooper[2,3], Rosa Fernández[1*]

[1] Metazoa Phylogenomics Lab, Biodiversity Program, Institute of Evolutionary Biology (CSIC - Universitat Pompeu Fabra). Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Spain
[2] Australian Centre for Evolutionary Biology and Biodiversity, Department of Ecology and Evolutionary Biology, School of Biological Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia
[3] Evolutionary Biology Unit, South Australian Museum, Adelaide, South Australia 5000, Australia
[4] Notranjska Museum Postojna, Kolodvorska c. 3, 6230 Postojna, Slovenia

*Corresponding authors: pau.balart@ibe.upf-csic.es, rosa.fernandez@ibe.upf-csic.es

ABSTRACT

Chemical cues in subterranean habitats differ highly from those on the surface due to the contrasting environmental conditions, such as absolute darkness, high humidity or food scarcity. Subterranean animals underwent changes to their sensory systems to facilitate the perception of essential stimuli for underground lifestyles. Despite representing unique systems to understand biological adaptation, the genomic basis of chemosensation across cave-dwelling species remains unexplored from a macroevolutionary perspective. Here, we explore the evolution of chemoreception in three beetle tribes that underwent at least six independent transitions to the underground through a phylogenomics spyglass. Our findings suggest that the chemosensory gene repertoire varies dramatically between species. Overall, no parallel changes in the net rate of evolution of chemosensory gene families were detected prior, during, or after the habitat shift among subterranean lineages. Contrarily, we found evidence of lineage-specific changes within surface and subterranean lineages. Our results reveal key duplications and losses shared between some of the lineages transitioning to the underground, including the loss of sugar receptors and gene duplications of the highly conserved ionotropic receptors IR25a and IR8a, involved in thermal and humidity sensing among other olfactory roles in insects. These duplications were detected both in independent subterranean lineages and their surface relatives, suggesting parallel evolution of these genes across lineages giving rise to cave-dwelling species. Overall, our results shed light on the genomic basis of chemoreception in subterranean beetles and contribute to deepen our understanding of the genomic underpinnings of adaptation to the subterranean lifestyle at a macroevolutionary scale.

INTRODUCTION

Chemoreception is an ancient biological function governing the perception of chemical information to mediate essential behavioral responses related to the detection of food, pathogens, predators, conspecifics, and optimal conditions for reproduction (Benton 2015). The essential role of chemoreception at multiple levels means the molecular components of this system are key candidates for important adaptive changes associated with life in new environments. Chemical stimuli are recognised by transmembrane proteins which transform volatile and soluble substances from the environment into electrical outputs through nerve impulses. These proteins are located on the dendritic membranes of sensory neurons housed in the sensilla of receptor organs, such as antennae and other sensory appendages along the body (Joseph and Carlson 2015). In arthropods, gene families involved in chemoreception have been studied in several lineages such as the insects (Sánchez-Gracia et al. 2009), decapods (Kozma et al. 2020), and spiders (Vizueta et al. 2017). This research has pinpointed six main protein families involved in chemoreception in insects: odorant receptors (ORs), gustatory receptors (GRs), ionotropic receptors (IRs), sensory neuron membrane proteins (SNMPs/CD36), chemosensory proteins (CSPs), and odorant-binding proteins (OBPs). GRs are the oldest animal chemoreceptor family, dating back to the origin of animals (Stocker 1994; Eyun et al. 2017). In protostomes, chemical detection is also mediated by IRs, which originated as a divergent clade of ionotropic glutamate receptors (IGluRs) and N-methyl-D-aspartate receptors (NMDARs) (Benton et al. 2009). IRs have a variety of functions, such as odorant perception, thermosensation, and hygrosensation, and are principally and widely expressed in peripheral sensory systems (van Giesen and Garrity 2017). IGluRs and NMDARs are glutamate-binding excitatory neurotransmitters that play a role in brain synaptic communication (Mayer and Armstrong 2004) and regulate developmental and reproductive processes in insects (Chiang et al. 2002) respectively. Furthermore, the sensory perception of smell in insects is mediated by ORs and SNMPs. ORs are a highly diverse chemoreceptor family that participate in the detection of airborne chemical compounds, particularly in appendages on the head (Clyne et al. 1999; Robertson et al. 2003; Thoma et al. 2019), and SNMPs play key roles in chemoreception such as a highly sensitive pheromone detection (Nichols and Vogt 2008; Gomez-Diaz et al. 2016). These transmembrane proteins are assisted by low-molecular-weight soluble proteins that act in parallel as carriers of hydrophobic compounds such as some odors and tastants, thus enhancing chemical perception. These carriers are coded by the CSPs and the insect-specific OBPs (Pelosi et al. 2014; Pelosi et al. 2018).

An extensive positive link between the extent of the chemosensory gene repertoire and the complexity of its chemical ecology has been found in insects and other arthropods occupying a large variety of ecological niches (Ngoc et al. 2016; Robertson et al. 2018; Andersson et al. 2019; Almudi et al. 2020). Chemosensory gene families have been observed

to experience a fast dynamic evolution, where gene gain and loss are the major source of variation, providing opportunities for adaptation to new habitats (Sánchez-Gracia et al. 2009).

Caves and other subterranean environments are characterized by a key selective pressure: the permanent absence of light. Moreover, underground habitats have a high climatic stability compared to surface habitats, with constant temperature and high environmental humidity (Lauritzen 2018). Furthermore, nutrient inputs tend to be heterogeneously distributed and are generally scarce (Kováč 2018). These conditions have impacted the sensory systems of its inhabitants at the phenotypic level. For example, the broadly convergent loss of eyes in cave-dwellers have been proposed to be compensated by enhanced mechanosensory and chemosensory capabilities (Poulson 1963). Sensory modifications include elongated antennae containing specific sensory organs in beetles (Accordi and Sbordoni 1977), modifications in the taste buds and lateral line in fishes (Varatharasan et al. 2009; Yoshizawa et al. 2014), and enhanced development of olfactory brain regions in crustaceans (Stegner et al. 2015). Nevertheless, very little is known about the molecular evolution of the chemosensory system in cave-dwelling fauna, particularly from a macroevolutionary perspective, with only a few studies so far focusing on single species at a time (Yang et al. 2016; Balart-García et al. 2021).

Cave beetles constitute an excellent system to understand the evolutionary processes leading to adaptation to subterranean habitats. Beetles comprise numerous instances of phylogenetically-distinct subterranean lineages with varying degrees of morphological change and ecological specialization related to their adaptation to underground habitats. The tribe Leptodirini (Leiodidae, Cholevinae) represents the largest animal radiation in subterranean habitats and is principally distributed in the North of the Mediterranean basin. This beetle tribe represents ancient colonizations of underground habitats (ca. 30 Mya) including multiple lineages with a varying degree of specialization to life in caves. Virtually all the species of the tribe Leptodirini are cave-dwelling or live in forest litter, being morphologically similar to cave-dwelling species (Ribera et al. 2010; Fresneda et al. 2011). Furthermore, diving beetles (Dytiscidae, Hydroporinae) also represent multiple independent instances of groundwater colonization in arid regions of Western Australia, specially in the tribes Bidessini and Hydroporini. These aquatic beetle tribes underwent underground colonization more recently (ca. 12–9 Mya) and show clear phenotypic differences between surface-dwelling and subterranean species (Leijs et al. 2012; Langille et al. 2021). The chemosensory gene repertoire has been characterized so far in a single cave beetle species, the Leptodirini beetle *Speonomus longicornis*. This species presents a diminished diversity of odorant and gustatory gene repertoires compared to polyphagous beetles inhabiting surface habitats, and a gene duplication of the ionotropic coreceptor IR25a, a highly conserved single-copy gene in protostomes (Balart-García et al. 2021) involved in thermal and humidity sensing among other

olfactory roles (Ni 2020). Nevertheless, the macroevolutionary processes governing chemosensory gene family evolution across cave-dwelling taxa have never been studied.

Here, we investigate the molecular evolution of the chemosensory system in beetles that have undergone independent instances of subterranean colonization using a genome-wide phylogenomic approach. Specifically, we annotated the main chemosensory gene families and characterized gene repertoire evolutionary dynamics in beetle species with a broad range of habitat preferences (i.e. aquatic, terrestrial, surface-dwelling and subterranean ecologies) to understand the role of gene gain, duplication and loss in the evolution of gene families involved in chemosensation. At least six underground transitions are represented by lineages of the tribes Leptodirini (terrestrial) and Bidessini and Hydroporini (aquatic), which allowed us to explore how the chemosensory gene repertoire was reshaped in the context of parallel subterranean evolution. Our study sheds light into the molecular evolution of the chemosensory systems of cave-dwelling fauna from a macroevolutionary perspective.

RESULTS AND DISCUSSION

Highly heterogeneous chemosensory gene repertoires in Coleoptera are potentially driven by diverse ecological preferences

Our data set includes 22 high quality transcriptomes for surface-dwelling and subterranean species of the tribes Bidessini, Hydroporini, and Leptodirini including at least six independent beetle lineages that colonized subterranean habitats. Moreover, we annotated candidate chemosensory genes for a total of 39 Coleoptera, 1 Strepsiptera and 1 Neuroptera species based on highly complete proteomes (Supplementary Table 1 and 2). Our results indicate a high variability in the extension and composition of the chemosensory gene repertoires across Coleoptera (Figure 1, Supplementary Table 3). ORs and OBPs are the most variable chemosensory gene families in terms of copy number variation. This result indicates that these two insect-specific chemosensory gene families, related to odorant perception, have undergone a more dynamic evolution compared to more ancient gene families that resulted in relatively less variable repertoires. Noticeably, species within Adephaga have more reduced chemosensory gene repertoires than the Polyphaga.

Large expansions in olfactory and gustatory related gene families may be linked to the perception of complex and heterogeneous environments in terms of chemical stimuli, which appear to be frequent in terrestrial ecologies (i.e., more prevalent in Polyphaga). Polyphaga species have more diverse feeding habits, including predators, scavengers, fungal grazers, and feeders of algae, bacteria and macrophytes. Conversely, the Adephaga species we studied are

mostly aquatic predators and they might not principally rely on airborne chemical clues to navigate in their environment. Furthermore, the early branching lineages of Polyphaga, represented by *Cyphon laevipennis* (Scirtidae), retain some ancestral habitat preferences (i.e. aquatic or semi aquatic larvae; Lawrence 2016) and show a reduced chemosensory repertoire, whereas other polyphagan species show remarkable expansions in specific gene families. The latter includes the firefly *Photinus pyralis* with 165 OBPs, and the red flour beetle and Asian longhorn beetle (*Tribolium castaneum* and *Anoplophora glabripennis*), which show extremely large gene repertoires that may reflect their extended chemosensory capabilities. Nevertheless, the only terrestrial adephagan species that we studied (i.e. *Calasoma frigidum*) shows a very reduced chemosensory gene repertoire and, in contrast, the aquatic polyphagan species (i.e *Hydrochus megaphallus*) shows a relatively large chemosensory gene repertoire. In essence, these results indicate that the extent of the chemosensory gene repertoires in Coleoptera might be strongly related to the habitat preferences and feeding habits, where aquatic predaceous beetles show relatively moderate chemosensory capabilities compared to terrestrial polyphagous beetles.

The sizes of the chemosensory repertoires in beetle tribes that have colonized underground habitats (i.e. Bidessini, Hydroporini and Leptodirini) do not show a clear pattern when comparing surface-dwelling and subterranean species. Remarkably, the subterranean species of Hydroporini and the subterranean Bidessini species of the genus *Limbodessus* show larger chemosensory gene repertoires than their corresponding surface-dwelling relatives within the same genera. These chemosensory gene repertoire expansions in subterranean species of the genus *Paroster* (Hydroporini) and *Limbodessus* (Bidessini) suggest that the subterranean environment might lead to chemosensory gene family expansions that therefore provide broader capabilities in the perception of chemical stimuli. This change could be related to the increased scavenging habits of subterranean species compared to the predatory preferences of their surface-dwelling relatives, or a dietary shift from predation to omnivory that broadly converges in many subterranean taxa (Gibert and Deharveng 2002, Hüppop 2012). Nonetheless, the other two genera of Bidessini (i.e. *Neobidessodes* and *Allodessus*) show the opposite pattern, in which the surface species have more chemosensory genes with a large expansion in the OBP family. Further comparisons, including surface-dwelling and subterranean relatives for each genus, need to be made to confirm the previously-mentioned pattern of enhanced chemosensory capabilities in subterranean aquatic beetles.

In the Leptodirini, we observe a contrasting pattern between two subterranean species: *Speonomus longicornis* has the largest chemosensory gene repertoire, whereas *Breuilia trianglum* exhibits a very reduced one. Moreover, closely related subterranean species such as *Leptodirus hochenwartii* and *Astagobius angustatus* show a high variation in their chemosensory repertoires, suggesting that these gene families have undergone extensive

changes at the lineage-specific level. Therefore, the chemosensory gene repertoire varies dramatically between species, and there seems to be no pattern with general trends when comparing surface to cave-dwelling species. Both surface-dwelling and subterranean species of the tribe Leptodirini are mainly detritivorous or saprophagous, feeding on carrion, fungi and biofilms. However, some cave-dwelling species show modifications of the mouthparts associated with a highly specific dietary niche. For instance, some subterranean species are adapted to a semiaquatic lifestyle, such as found in hygropetric habitats on cave walls, and filter organic particles with their modified mouthpieces (Moldovan 2004). Lineage-specific dietary specializations in Leptoridini, may have contributed to the highly diverse chemosensory gene repertoires that we observed in surface and cave-dwelling species.

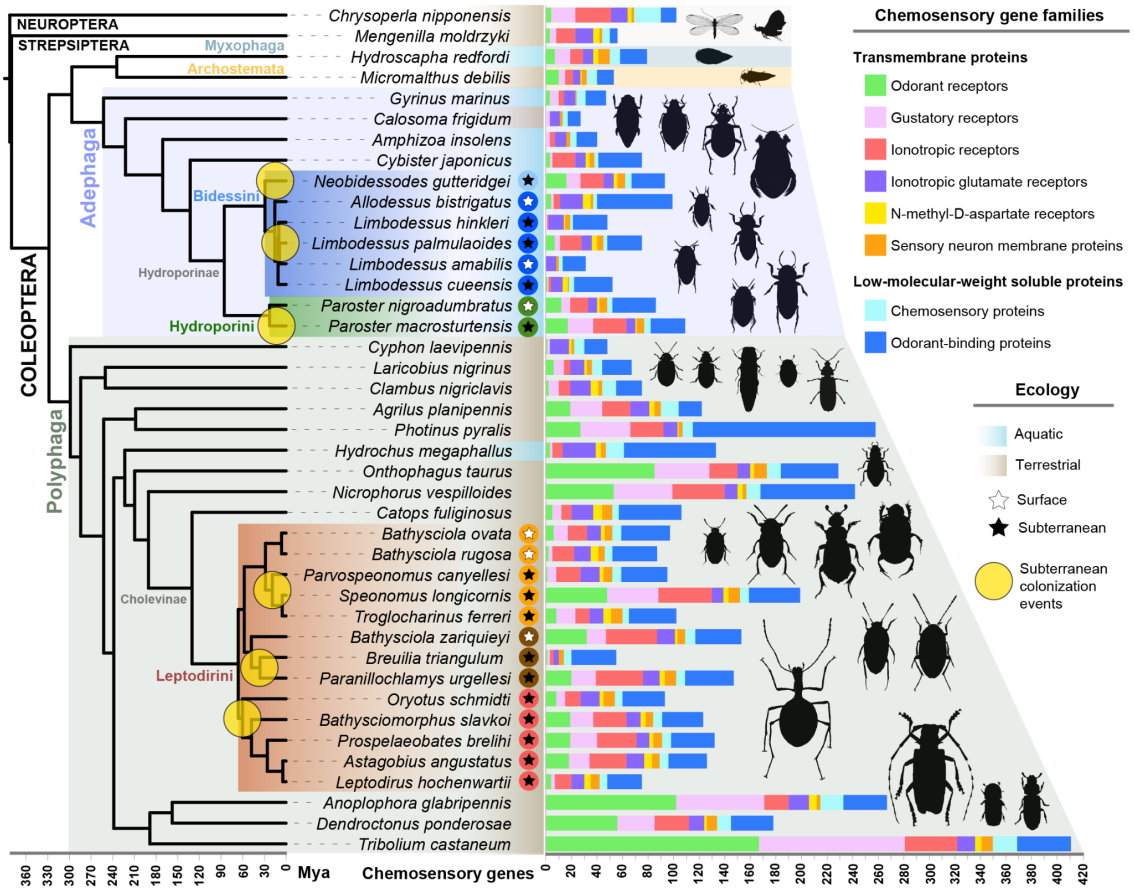

Figure 1. Phylogenetic context and size of the chemosensory gene repertoires in Coleoptera. Time-calibrated phylogeny of the studied Coleoptera and outgroups, including six independent subterranean colonization events in the tribes Bidessini, Hydroporini and Leptodirini. The central column indicates the ecology of the studied species (i.e. aquatic - terrestrial, surface - subterranean). Bar plots indicate the total number of chemosensory genes annotated for each species.

Rapid and complex evolution of the chemosensory system in subterranean beetle lineages

We used maximum likelihood-based phylogenetic comparative methods to explore parallel evolution in the chemosensory gene repertoires of subterranean lineages. We investigated significant gene gain and loss shifts for each chemosensory gene family (i.e. the difference between gene gain and loss rates, net rate hereafter) prior, during, and after the subterranean transition. To do this, we compared a global rates (G) estimation model (i.e. assuming the same net rate for all the branches) to several branch-specific rate models. These branch-specific models assumed net rate differences in the branches of the most recent common ancestor (MCRA) of each beetle tribe (i.e. Bidessini, Hydroporini and Leptodirini), in the branches where the underground habitat shift (HS) occurred or in all the subterranean branches (SB) compared to the background. We performed these analyses including all taxa and also at a lineage-specific level, with only the tribe Leptodirini and the tribes Bidessini + Hydroporini separately, which respectively represent terrestrial and aquatic subterranean transitions. In the Leptodirini-specific analysis, we added an extra branch-specific model that tested significant differences in the net rate of chemosensory gene families in two branches leading to highly modified lineages (HML) of cave beetles (i.e. *Speonomus longicornis* - *Troglocharinus ferreri*; *Astagobius angustatus* - *Leptodirus hochenwartii*). These two lineages independently developed contracted life cycles (i.e. single larval instar development) and remarkable anatomical changes (i.e. extreme enlargement of the body shape and appendages) associated with a higher degree of subterranean specialization compared to their cave-dwelling relatives (Deleurance-Glaçon 1963; Delay 1978; Cieslak et al. 2014).

None of the branch-specific rate models were significantly better-fitting compared to the global rate model for any chemosensory gene family, indicating no parallel changes in the net rate of evolution of chemosensory gene families occurred prior, during, or after the subterranean transition in the selected tribes (Supplementary Table 4). Moreover, we could not obtain the net rate estimation for the odorant receptors (OR) and gustatory receptors (GR) under any branch-specific rates model due to the high heterogeneity of the gene counts across the selected taxa and the complexity of the branch-specific models. At the lineage-specific level, we obtained the net rate estimations under all the models and it was possible to estimate the rate of evolution of ORs. In the Leptodirini-specific analysis, ORs resulted in a significantly better fit under the habitat shift (HS) model, but only after excluding the outlier *S. longicornis.* Our results thus indicate a contraction of the OR repertoire when lineages transitioned underground in Leptodirini. Nonetheless, given the outlier and the limited representation of surface-dwelling species, the evolution of odorant receptors in cave-dwelling fauna need to be further investigated to obtain more robust evidence.

We also investigated the estimated gene gains and losses per branch for each chemosensory gene family using the results of the global rates model. This allowed us to explore in detail where the most relevant pulses of chemosensory gene repertoire change occurred in the phylogenetic context of this study (Figure 2). Our results indicate that gene family expansion was more relevant than contraction in most of the species. Nevertheless, we found some species that primarily show gene family contractions (i.e. *Limbodessus amabilis*, *Troglocharinus ferreri*, *Breuilia triangulum* and *Leptodirus hochenwartii).* Despite not being significant under the previously mentioned MCRA branch-specific model, some gene families experienced parallel changes in the MRCA of each tribe. For instance, OBPs and CSPs experienced parallel expansions, and IGluRs were parallelly contracted in the MRCA of the three tribes. On the other hand, the IR family was contracted in the MRCA of Leptodirini and Bidessini, and expanded in Hydroporini. The SNMP family was only expanded in the MRCA of Leptodirini and Hydroporini. These parallel and exclusive changes could represent examples of genomic exaptation that potentially led to chemosensory adaptation in underground habitats. Moreover, we did not observe substantial changes in the branches where species transitioned underground, suggesting that the chemosensory families experienced lineage-specific changes within surface and subterranean lineages. More remarkably, and despite not being significant in the HML model for the Leptodirini analysis, the ancestral estimations of gene gain and loss across the phylogeny suggested a parallel reshaping of the chemosensory repertoire in the highly modified lineages of Leptodirini, characterized by a remarkable parallel expansion of GRs, IRs, and, to a lesser extent, an expansion of NMDARs. These results suggest highly specialized subterranean lineages that colonized caves >30 Mya could have experienced key parallel changes in their chemosensory repertoires. When further exploring the expansions and contractions in these highly modified species we observed contrasting patterns, in which *A. angustatus* and *S. longicornis* show large expansions of GRs and IRs compared to the gene family dynamics of *L. hochenwartii* and *T. ferreri,* that essentially consist of contractions. Likewise, the lineage leading to the surface-dwelling species *Bathysciola ovata* + *Bathysciola rugosa* also experienced a relevant expansion in several chemosensory gene families. Altogether, these results indicate that chemosensory gene families experienced rapid evolution including expansion and contractions in surface and subterranean species. Contrasting patterns in closely related species suggest that no significant quantitative changes occurred due to subterranean colonization, and thus that changes occurred independently within subterranean ecologies.

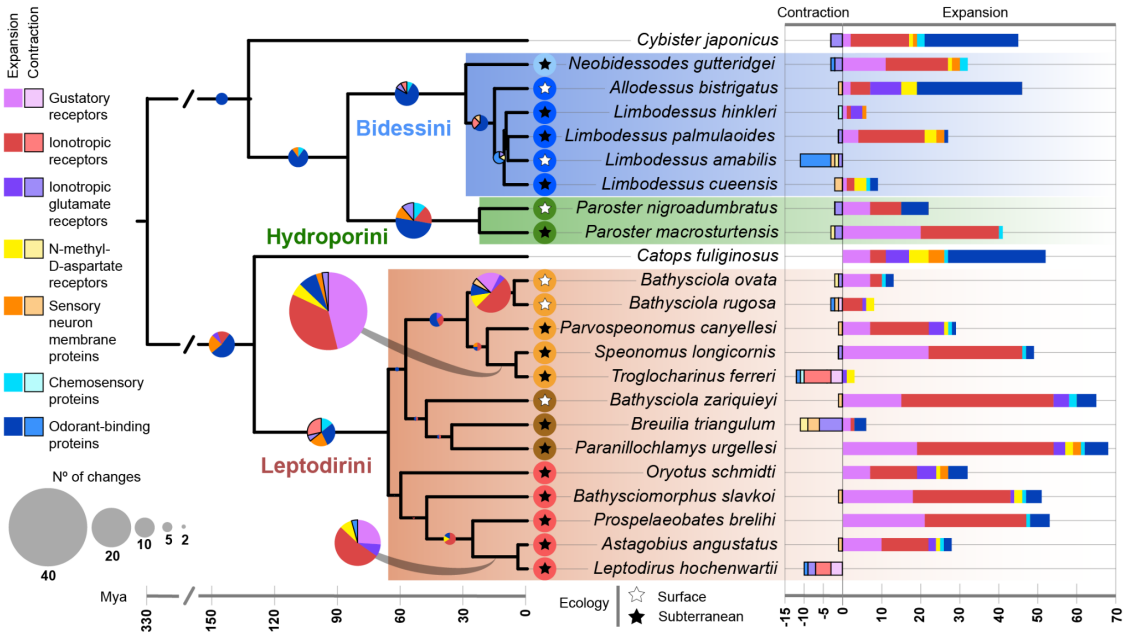

Figure 2. Ancestral contractions and expansions of the chemosensory gene families. Time-calibrated tree indicating the total number of gene gains and losses of each chemosensory gene family in each branch. Pie charts are scaled to the total number of gene gains and losses of each chemosensory gene family. Bar plots indicate the total number of gene losses (left) and gene gains (right) in the terminal branches. Coloured ranges indicate the tribes Bidessini, Hydroporini and Leptodirini and coloured circles indicate the independent instances of underground colonization.

The odorant receptor (OR) gene family was significantly contracted in the subterranean lineages of Leptodirini, with the exception of *S. longicornis* that showed the most extended chemosensory gene repertoire within Leptodirini (Figure 3). Furthermore, our analyses revealed a low number of genes in this gene family in the MRCA of each clade and a progressive expansion towards the tips, suggesting that a fast dynamic of expansions occurred at a lineage-specific level and was the driving process guiding the evolution of ORs.

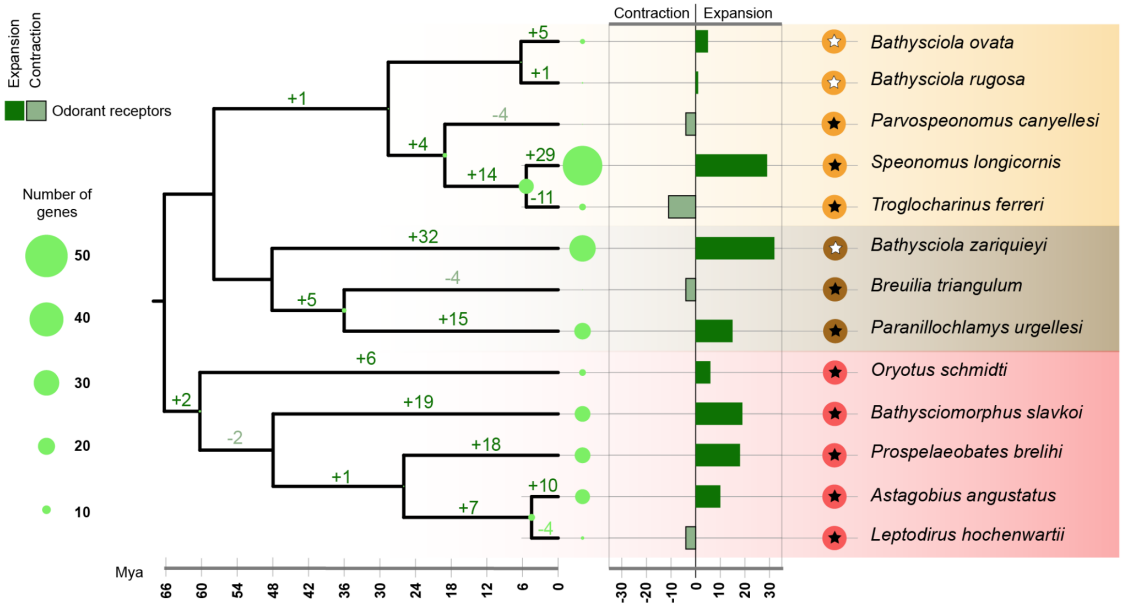

Figure 3. Odorant receptors (ORs) evolution in the tribe Leptodirini. Time-calibrated phylogeny indicating the estimated size of the OR repertoires (green dots) and the number of gene gains and losses. Barplots indicate the proportion of gene gain and loss in the terminal branches. White stars indicate surface-dwelling species, black stars indicate subterranean species and coloured circles indicate lineages representing independent underground colonization events.

Parallel gene gains and losses in key chemosensory gene families may have facilitated the adaptation to life in caves

We performed maximum-likelihood phylogenetic inferences for each chemosensory gene family in order to (i) validate the annotations, (ii) identify conserved chemosensory gene clades in insects and (iii) explore their relationships among Coleoptera. Most of the conserved chemosensory clades appear highly supported in our phylogenetic inferences (Figure 4; see also Supplementary Figures 1-7). However, there are some differences in the relationships between some clades compared to those obtained in previous studies that included a more reduced set of Coleoptera species. For instance, the OR2A clade was not a monophyletic group in our analysis and members of this clade, previously reported in *T. castaneum,* were placed into two supported subclades (Figure 4). Moreover, we detected nine highly-supported CSP clades for Coleoptera (i.e. A-I) and one composed only by Neuroptera and Strepsiptera sequences (i.e. O). Some of these CSP clades correspond to lineage-specific expansions, for instance the clade I is specific to *Tribolium castaneum*, H is specific to Tenebrionoidea + Chrysomeloidea + Curculionoidea and E is specific to Hydrophiloidea + Scarabaeoidea + Staphylinoidea (Supplementary Figure 1). Regarding the OBP clades, we found two supported clades for the antennal binding proteins II (ABPII) and for Classic OBPs and four clades of

Minus-C OBPs. Interestingly, most of the OBPs of *Photinus pyralis*, which would represent one of the largest OBP repertoires reported in insects, correspond to divergent OBPs with multiple duplications (Supplementary Figure 2). Nonetheless, we were not able to characterize any OR, GR and some of the conserved IR clades in some species (Supplementary Table 3), indicating that a more exhaustive approach needs to be used to clarify the evolution of these transmembrane chemoreceptor gene families in Coleoptera. Gene expression data from different chemosensory appendages showing functional specificity, such antennae and mouthparts, and comparisons with highly complete genomes of the same species would also help to characterize in detail these chemoreceptor families.

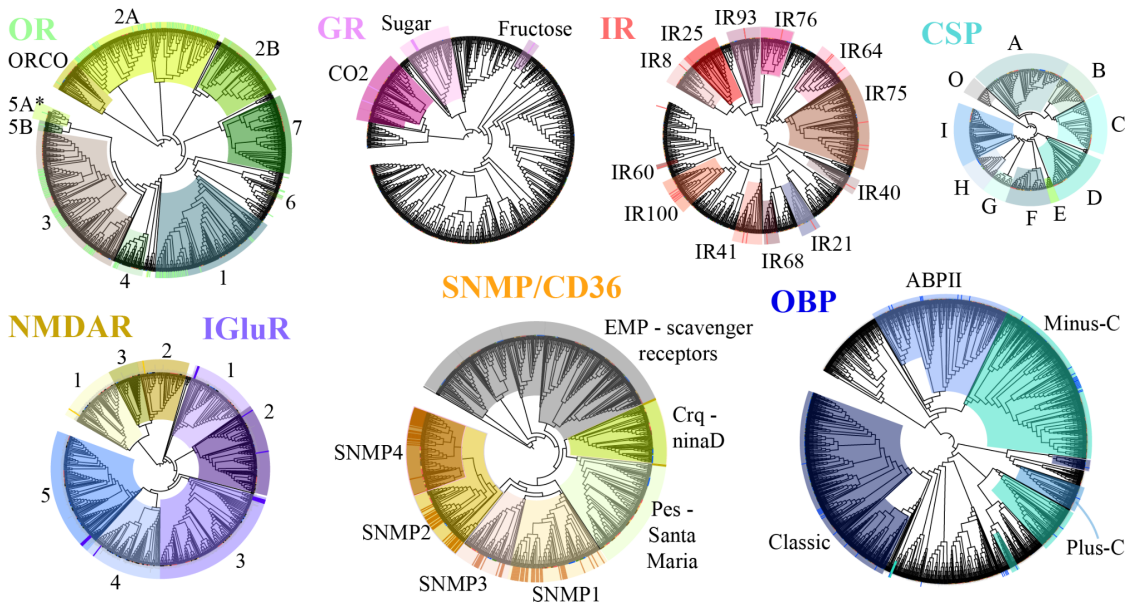

Figure 4: Chemosensory gene families and supported clades in Coleoptera. Maximum-likelihood phylogenetic trees for each chemosensory gene family as follows: odorant receptors (OR), gustatory receptors (GR), ionotropic receptors (IR), N-methyl-D-aspartate receptors (NMDAR), ionotropic glutamate receptors (IGluR), sensory neuron membrane proteins (SNMP/CD36), chemosensory proteins (CSP) and odorant-binding proteins (OBP). Coloured ranges indicate significantly supported clades (i.e. bootstrap >75%) identified with reference sequences from a variety of studies (i.e. indicated with darker outer strips, see Methods). OR5A (*) genes of *Tribolium castaneum* are collapsed (i.e. 227 ORs). See Supplementary Figures 1-7 for more detailed information.

Remarkably, the conserved gustatory receptors related to the perception of sugar and fructose were not detected in any Leptodirini species nor in *Catops fuliginous* (both in the same subfamily, Cholevinae) (Figure 5A). However, *Nicrophorus vespilloides* (same superfamily, Staphylinoidea) showed three gene copies of candidate sugar receptors. Our results thus indicate that the gustatory perception of sugar and fructose has potentially been lost in

Cholevinae. On the other hand, we detected sugar receptors in two subterranean species of the tribes Hydroporini and Bidessini (*Paroster macrosturtensis* and *Limbodessus palmulaoides*), indicating the aquatic clade retains sugar perception capabilities. Furthermore, the apparently conserved fructose receptor clade previously characterized in some phytophagous Polyphaga species is also present in *Laricobius nigrinus, a* predatory specialist of the hemlock woolly adelgid, indicating that this fructose receptor clade is not exclusive to species merely feeding on plants, but is also conserved in species whose diet may include substantial inputs of fructose (Cohen and Cheah 2015).

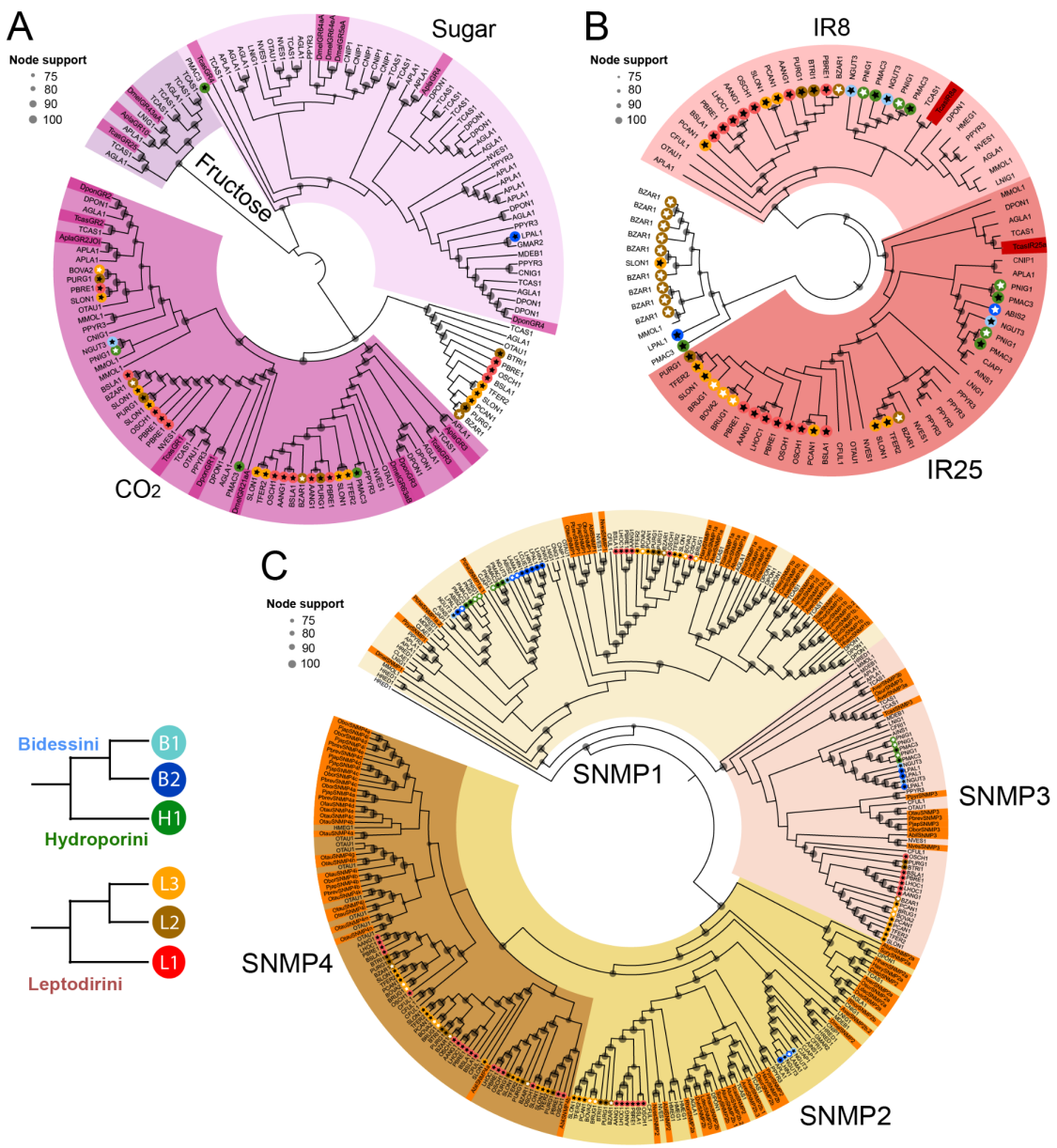

Figure 5. Conserved gustatory (GR) and ionotropic (IR) receptor clades and sensory neuron membrane protein (SNMP) clades in Coleoptera. Pruned maximum-likelihood gene trees of GR (A), IR (B) and SNMP (C). White stars represent surface-dwelling species and black stars

subterranean species. The dark outer strips correspond to reference genes used to identify the main clades (see Methods). For the complete gene trees see Supplementary Figures 4, 5 and 6.

We further explored the conserved clade of CO_2 gustatory receptors, which typically consist of two or three gene copies in Coleoptera (Andersson et al. 2019). Two subterranean species of the tribe Leptodirini (i.e. *Speonomus longicornis* and *Prospelaeobates brelihi*) showed more than three copies of these highly conserved CO_2 receptors (Figure 5A, Supplementary Figure 4), suggesting parallel gene duplications of these CO_2 receptors in independent subterranean lineages. Remaining Leptodirini species showed a variable number of CO_2 receptors (from zero to three). On the other hand, very few CO_2 receptors were identified in Hydroporini (i.e. one copy in the surface-dwelling species and two copies in the subterranean species), and just one copy in a surface-dwelling species of Bidessini. Moreover, we detected a highly supported sister clade of these CO_2 receptors consisting of sequences of primarily Leptodirini species, but also other Polyphaga species. In summary, these results indicate that CO_2 receptors are less conserved than originally suggested in previous studies and that possibly the species with absent receptors may perceive CO_2 via other kinds of receptors.

Ionotropic receptors (IR), highly conserved in virtually all protostomes, also showed some novelties in the studied beetle tribes. A surface-dwelling species and several subterranean species of the tribe Leptodirini (*Bathysciola rugosa*, *Speonomus longicornis*, *Troglocharinus ferreri*, *Prospelaeobates brelihi*, *Oryotus schmidti*) and Hydroporini (*Paroster macrosturtensis* and *Paroster nigroadumbratus*) showed candidate gene duplications of the IR25a gene, suggesting a parallel duplication of this gene not only in independent subterranean lineages, but also in their surface relatives (Figure 5B, Supplementary Figure 5). Interestingly, the IR25a gene has been shown to be related to temperature and humidity sensing in *D. melanogaster* among other olfactory roles (Knecht et al. 2017) potentially suggesting this candidate gene duplication enhanced the perception of key environmental variables in surface and subterranean species of both Leptodirini and Hydroporini. Despite the Hydroporini species being aquatic and inhabiting environments with a highly constant temperature, *Paroster nigroadumbratus* inhabit temporary waters that are impacted by droughts and floods on the surface (Watts and Leijs 2008), thus a fine hygrosensory and thermosensory system might represent a clear advantage to detect seasonality. Surface-dwelling and subterranean species of the Leptodirini tribe inhabit environments with a high environmental humidity and constant temperature and they show a low desiccation resistance, thus detecting small variations in temperature and humidity may facilitate them to navigate and find the optimal conditions within their particular habitat. The gene duplication of IR25a has been described previously in *S. longicornis* (Balart-García et al. 2021), and in this study we expand the observation of several cave beetle species and their surface relatives. In addition, a candidate duplication was also

detected in the IR8a (i.e. a basic co-receptor for odorant reception of multiple compounds together with other divergent IRs) for both surface-dwelling and cave species of Hydroporini and for two subterranean species of Leptodirini (*A. angustatus* and *P. brelihi*). A duplication of IR8a was only previously observed in copepods (Eyun et al. 2017), indicating that these cases in Leptodirini and Hydroporini species represent the first occurrence in insects.

We also detected several copies of the recently described beetle-specific clade SNMP4 found in some scarabaeid and staphylinid species (Zhao et al. 2020) (Figure 5C, Supplementary Figure 6). Our results revealed the presence of SNMP4 in *Hydrochus megaphallus* (Hydrophyloidea) and multiple copies were detected in species of the tribe Leptodirini and in *Catops fuliginosus,* the latter clustering together with the genes of their closest relative within Staphylinoidea (*Aleochara bilineata*). This finding indicates that the SNMP4, which essentially consists of a SNMP2 subclade, is specific to Hydrophiloidea, Scarabaeoidea and Staphylinoidea.

Here, we show that the evolution of the chemosensory gene repertoire is highly dynamic across cave beetle lineages, with no common pattern of variation among lineages representing independent subterranean colonizations. Our results indicate gene gain and expansion were major forces driving the evolution of gene families involved in chemoreception, suggesting that the chemosensory system increased in complexity in subterranean lineages, probably to compensate for the loss of information through the visual system. Key specific gene duplications and losses occurred prior to underground transitions, epitomized by the duplication of the genes IR25a and IR8a (involved in thermal and humidity sensing among other olfactory roles) in cave lineages and their surface sister species, or the loss of sugar and fructose receptors prior to the diversification of the tribe Leptodirini that comprises several cave lineages. All in all, our findings pave the way to a deeper comprehension of macroevolutionary trends that underlie adaptation to life underground.

MATERIALS AND METHODS

Data collection

We generated data for a total of 21 Coleoptera species that inhabit aquatic and terrestrial habitats both including surface-dwelling and subterranean ecologies. The aquatic lineage (Dytiscidae, Hydroporinae) include two species from the tribe Hydroporini and six species from the tribe Bidessini. The terrestrial lineage (Leiodidae, Cholevinae), include a total of 13 species from the tribe Leptodirini and one species from the tribe Catopini (i.e. *Catops fuliginosus*, representing a closely related outgroup). We obtained the transcriptome of one of

these 13 Leptodirini species from Balart-Garcia et al. (2021) (i.e. *Speonomus longicornis*). The terrestrial species inhabiting surface-dwelling habitats were collected in forest-litter using an entomological litter reducer and a metallic sieve of 3mm aperture. Those inhabiting subterranean habitats were captured manually in caves using an entomological aspirator. The surface-dwelling aquatic species were captured in shallow still water and temporary pools using aquatic nets whereas for the subterranean species we accessed their habitat though groundwater-monitoring boreholes and we captured specimens with a mesh net attached to a fishing rod. When the localities were remote, we kept specimens in RNA-later and cold conditions to prevent degradation. When possible, specimens were kept alive in thermo boxes. For the terrestrial species we used moss to maintain the humidity. The aquatic species were transported in tubes containing the same water of the sampling site. Samples were transported in dark and constant temperature, simulating natural conditions. Specimens were identified and flash-frozen with liquid nitrogen, then samples were stored at -80ºC. Total-RNA was extracted in cold, sterile and RNAse free conditions using different protocols, including (i) the selective Phenol/Chloroform extraction following a lysis through guanidinium thiocyanate buffer as described in (Sambrook et al. 1989) and modified in Balart-Garcia et al. (2021), (ii) the HigherPurityTM Tissue Total RNA Extraction Kit (Canvax Biotech) protocol and (iii) Qiagen Allprep RNA/DNA Isolation Kit. For the libraries of a single specimen of the aquatic lineage we used Nugen kit, the multiple individual samples were prepared with the TruSeq stranded mRNA libraries at the sequencing facilities (see below). A quality control check was performed using Agilent Technologies 2100 Bioanalyzer using the RNA Integrity Number (RIN). Libraries were sequenced in paired-end mode on the Illumina platforms HiSeq2500, and Novaseq 6000. *Limbodessus amabilis* data were obtained through an Iso-Seq protocol. A PacBio SMRTBell Template Prep Kit 1.0 SPv3, 1 “Non sized” selected library and 1 Size Selected library (4.5-10kb) per sample were used. Libraries were sequenced in a PacBio Sequel Isoseq platform. mRNA sequencing was performed at the Australian Genome Resource Facility (AGRF, Australia), at the National Center of Genomic Analysis (CNAG, Spain) and at Macrogen (South Korea). For more detailed information about species, samples, localities, laboratory protocols, sequencing details see Supplementary Table 1.

In order to study the evolution of chemosensory gene families across Coleoptera, we collected public genomic and transcriptomic data from a further 17 Coleoptera species and two outgroups (i.e. Strepsiptera and Neuroptera). The criteria for selecting these species consisted of checking (i) their phylogenetic relationships (i.e. following the topology of the Coleoptera tree of life described in (McKenna et al. 2019), (ii) the availability of the raw data in public repositories and (iii) the quality of the data based on the completeness score obtained after processing the raw data (see below). For more detailed information about species and public data repositories see Supplementary Table 2.

Data preparation

We followed the raw data processing pipeline explained in (Fernández et al. 2022) with a few modifications. For RNA-seq data, adapters and low-quality reads were trimmed using fastp v.0.20.041, and de novo assembly was performed using Trinity v2.1142 with default parameters. Candidate coding regions were predicted using TransDecoder v5.5.0 (https://github.com/TransDecoder/TransDecoder) and the contaminant sequences were filtered with BlobTools2 (Challis et al. 2020) and a previous detection of all non-metazoan sequences through a DIAMOND (Buchfink et al. 2021) BLASTP search (--sensitive --max-target-seqs 1 --evalue 1e-10) against the NCBI non-redundant (NCBI NR) protein sequence database (http://www.ncbi.nlm.nih.gov/). Finally, longest isoforms of each transcript were retained as final candidate coding regions for further analyses. For the PacBio data obtained for *Limbodessus amabilis* we used Iso-Seq v.3.1.2 (https://github.com/PacificBiosciences/IsoSeq) to process the raw reads, ANGEL (https://github.com/PacificBiosciences/ANGEL) to predict open reading frames, and CD-HIT v.4.8.1 (Fu et al. 2012) to cluster highly identical sequences with a similarity threshold of 0.9. For whole-genome sequencing data, we downloaded the predicted proteome for published genomes. We assessed the quality and completeness of newly assembled transcriptomes and the downloaded proteomes using BUSCO v4.1.4 (Manni et al. 2021) in protein mode, obtaining an estimation of the completeness of the gene content based on the sets of benchmarking universal single-copy orthologs of the Insecta database (insecta_odb10). The final genomic data set resulted in gene completeness score percentage higher than 80% (i.e. considering the sum of the percentages of both the complete and fragmented genes; mean 93.15% BUSCO completeness), with the exception of the newly generated transcriptomes of *Limbodessus cueensis* and *L. hinkleri* that resulted in a BUSCO completeness of 73 and 76% respectively (see Supplementary Table 1 and 2 for further details). We finally obtained a total of 41 proteome sets.

Chemosensory families characterization and phylogenetic inferences

We annotated genes encoding chemosensory related proteins (OR, GR, IR, IGLUR, NMDAR, SNMP, OBP and CSP) for the 41 proteomes by using BITACORA v.1.3 (Vizueta et al. 2020) in “protein mode”. This sequence similarity based software combines BLAST and HMMER searches using custom databases that contain sequences of the gene families of interest. We generated curated databases for each chemosensory gene family which include sequences of *Drosophila melanogaster* (flybase.org), that had been experimentally confirmed (i.e. GR, IGluR, NMDARs and SNMP/CD36), and sequences of several Coleoptera species that were characterized in previous studies using comparative genomics methods (OR, GR, IR, SNMP, CSP, OBP) (Dippel et al. 2014, 2016; Schoville et al. 2018; Andersson et al. 2019; Zhao et al. 2020). Due to the high heterogeneity of some chemosensory gene families and the

inclusion of partial genes in our proteome sets, the dubious hits obtained with each database (e.g. result significant both for OR and GR hits or IGluR and IR hits) were identified and manually validated using additional protein searches with InterPro (Blum et al. 2021). Proteomes of species whose chemosensory gene families were characterized in previous studies (i.e. *Tribolium castaneum*, *Agrilus planipennis*, *Onthophagus taurus*, *Nicrophorus vespilloides*, *Anoplophora glabripennis* and *Dendroctonus ponderosae*) were re-annotated with our approach in order to follow the same methodology for all the data sets.

We conducted maximum-likelihood phylogenetic inferences for each chemosensory gene family in order (i) to validate the annotation of the candidate chemosensory gene sets and discard dubious sequences, (ii) to identify highly conserved chemosensory subclades across insects, and (iii) to study their phylogenetic relationships across beetles. To do this, we included some sequences of the curated databases in the phylogenetic inferences as references for the main chemosensory gene family clades. These references consisted of GR, IGluR, NMDAR, SNMP/CD36 sequences of *D. melanogaster*, OR, GR, IR, CSP, SNMP/CD36 and OBP sequences of *T. castaneum,* GR sequences of *A. planipennis* and *D. ponderosae* and a variety of Coleopteran SNMP sequences obtained from Zhao et al. (2020). We aligned the amino acid sequences for each chemosensory gene family using MAFFT v.7.407 (Katoh and Standley 2013) with a maximum of 1,000 iterations. We trimmed the alignments with trimAl v.1.4.1 (Capella-Gutiérrez et al. 2009) using the option -gt 0.4 and manually discarded sequences with less than 10% of the total amino acid positions in the alignment. We used FastTree v.2.1.11 under the LG model (Price et al. 2010) to generate a gene tree that was then implemented as a guide tree for the final phylogenetic inference with IQ-TREE v.2.1.3 (Minh et al. 2020) using the mixture model LG+C20+F+G with the site-specific posterior mean frequency model (PMSF; Wang et al. 2018) and the ultra-fast bootstrap option (Hoang et al. 2018). We used iTOL v6 to visualize and render the phylogenetic trees (Letunic and Bork 2021).

Molecular dating

We used the time calibrated tree that was generated in Balart-Garcia et al. (submitted) following an orthology based approach. See details in Supplementary Methods.

Estimation of the evolutionary dynamics of the chemosensory gene repertoire

We used BadiRate (Librado et al. 2012) to study the evolutionary dynamics of the chemosensory gene families under the birth, death and innovation model (BDI). This software uses a maximum-likelihood based approach to estimate gene gains and losses across all branches in a given phylogenetic context. We performed this estimation for each chemosensory gene family by using the candidate gene counts (Supplementary Table 3) and the

time-calibrated tree. Furthermore, to investigate the most relevant evolutionary events that could lead to parallel significant changes in the chemosensory gene families, we examined significant changes in the net rate (i.e. the difference between gene gain and loss rates, henceforth 'net rate') under different branch models. We compared a "global rates" (G) estimation model—which assumes that all of the branches have the same net rate—to several "branch-specific" rates models, that assume that a certain set of selected branches have different net rates compared to the background. These branch-specific models consisted of (i) the branches of the most recent common ancestors of the tribes (MCRA model), (ii) the branches leading to the underground habitat shift (HS model) and (iii) all the branches of subterranean lineages together (SB model). These analysis were performed including all the taxa, but in order to explore lineage-specific shifts in the chemosensory gene families, we performed the analysis at a narrower phylogenetic context either with only the aquatic tribes (i.e. Bidessini and Hydroporini) or the terrestrial tribe (i.e. Leptodirini) separately. For the terrestrial tribe analysis we performed an extra branch-specific model which assumes a rate shift in the branches of the highly modified lineages (HML) compared to the background and thus allowing us to explore whether remarkable changes in the chemosensory repertoire occurred within a subterranean ecology. These highly modified lineages correspond to two clades of the tribe Leptodirini that parallelly developed a life cycle contraction and extremely modified body shape and elongation of sensory appendages (i.e. *Speonomus longicornis* - *Troglocharinus ferreri*; *Astagobius angustatus* - *Leptodirus hochenwartii*) (Deleurance-Glaçon 1963; Delay 1978; Cieslak et al. 2014). Finally we calculated the akaike information criterion (AIC) and compared the global rates model to each branch-specific model independently at every phylogenetic level (i.e. all taxa, aquatic lineage, terrestrial lineage) in order to find the best fitting model.

DATA AVAILABILITY STATEMENT

Raw reads have been deposited in the National Center for Biotechnology Information (NCBI; BioProject accession number PRJNA902350. All processed data and custom scripts have been deposited in github (https://github.com/MetazoaPhylogenomicsLab/Balart_Garcia_et_al_chemosensory_gene_repertoire_evolution_subterranean_coleoptera).

ACKNOWLEDGEMENTS

PBG was supported by an FPI grant (grant agreement no. BES-2017-081050) financed by MCIN/AEI /10.13039/501100011033 and by European Social Fund (ESF) 'Investing in your Future' and the Systematics Research Fund 2020 (Linnean Society of London and the

Systematics Association). RF acknowledges support from the following sources of funding: Ramón y Cajal fellowship (grant agreement no. RYC-2017-22492 funded by MCIN/AEI /10.13039/501100011033 and ESF 'Investing in your future'), project PID2019-108824GA-I00 funded by MCIN/AEI/10.13039/501100011033, and by the European Research Council (ERC) under the European's Union's Horizon 2020 research and innovation programme (grant agreement no. 948281)). IR acknowledges support from project CGL2016-76705-P funded by MCIN/ AEI /10.13039/501100011033/ and by ERDF 'A way of making Europe'. SJBC acknowledges support from the Australian Research Council (grant agreement DP18010385 and DP120102132). We also thank Centro de Supercomputación de Galicia (CESGA) for access to computer resources.

CONTRIBUTIONS

PBG, IR and RF designed and conceived the research. PBG, SP, IR and SJBC collected samples and provided valuable knowledge about the studied systems. TMB and PGBH generated genomic data of the aquatic lineages. PBG generated genomic data of the terrestrial lineages, processed both newly generated and public genomic data and performed all analyses. PBG prepared the figures and other provided materials. PBG, SJBC and RF interpreted and discussed the results. IR, SJBC and RF provided resources. RF supervised the study. PBG and RF drafted the paper. All authors contributed to the final version of the manuscript.

REFERENCES

Accordi FS, Sbordoni V. 1977. The Fine Structure of Hamann's Organ in Leptodirus Hohenwarti, a Highly Specialized Cave Bathysciinae: (Coleoptera, Catopidae). *Int. J. Speleol.* 9:153.165.

Almudi I, Vizueta J, Wyatt CDR, de Mendoza A, Marlétaz F, Firbas PN, Feuda R, Masiero G, Medina P, Alcaina-Caro A, et al. 2020. Genomic adaptations to aquatic and aerial life in mayflies and the origin of insect wings. *Nat. Commun.* 11:2631.

Andersson MN, Keeling CI, Mitchell RF. 2019. Genomic content of chemosensory genes correlates with host range in wood-boring beetles (Dendroctonus ponderosae, Agrilus planipennis, and Anoplophora glabripennis). *BMC Genomics.* 20.

Balart-García P, Cieslak A, Escuer P, Rozas J, Ribera I, Fernández R. 2021. Smelling in the dark: Phylogenomic insights into the chemosensory system of a subterranean beetle. *Mol. Ecol.* 30:2573–2590.

Benton R. 2015. Multigene Family Evolution: Perspectives from Insect Chemoreceptors. *Trends Ecol. Evol.* 30:590–600.

Benton R, Vannice KS, Gomez-Diaz C, Vosshall LB. 2009. Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. *Cell* 136:149–162.

Blum M, Chang H-Y, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, Nuka G, Paysan-Lafosse T, Qureshi M, Raj S, et al. 2021. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res.* 49:D344–D354.

Buchfink B, Reuter K, Drost H-G. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* 18:366–368.

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

Challis R, Richards E, Rajan J, Cochrane G, Blaxter M. 2020. BlobToolKit - Interactive Quality Assessment of Genome Assemblies. *G3* 10:1361–1374.

Chiang A-S, Lin W-Y, Liu H-P, Pszczolkowski MA, Fu T-F, Chiu S-L, Holbrook GL. 2002. Insect NMDA receptors mediate juvenile hormone biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 99:37–42.

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*. 281:20132978.

Clyne PJ, Warr CG, Freeman MR, Lessing D, Kim J, Carlson JR. 1999. A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in Drosophila. *Neuron* 22:327–338.

Cohen AC, Cheah CA. 2015. Interim diets for specialist predators of hemlock woolly adelgids. *Entomology, Ornithology & Herpetology* 4:1.

Delay B. 1978. Milieu souterrain et écophysiologie de la reproduction et du développement des Coléoptères Bathysciinae hypogés. *Mém. Biospéol.* 5:1–349.

Deleurance-Glaçon S. 1963. Recherches sur les Coléoptères troglobies de la sous-famille Bathysciinae. *Annu Sci Nat Zool Paris* 12:1–172.

Dippel S, Oberhofer G, Kahnt J, Gerischer L, Opitz L, Schachtner J, et al. 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* 15:1141

Dippel S, Kollmann M, Oberhofer G, Montino A, Knoll C, Krala M, Rexer K-H, Frank S, Kumpf R, Schachtner J, et al. 2016. Morphological and Transcriptomic Analysis of a Beetle Chemosensory System Reveals a Gnathal Olfactory Center. *BMC Biol.* 14:90.

Eyun S-I, Soh HY, Posavi M, Munro JB, Hughes DST, Murali SC, Qu J, Dugan S, Lee SL, Chao H, et al. 2017. Evolutionary History of Chemosensory-Related Gene Families across the Arthropoda. *Mol. Biol. Evol.* 34:1838–1862.

Fernández R, Tonzo V, Guerrero CS, Lozano-Fernandez J, Martínez-Redondo GI, Balart-García P, Aristide L, Eleftheriadi K, Vargas-Chávez C. 2022. MATEdb, a data repository of high-quality metazoan transcriptome assemblies to accelerate phylogenomic studies. *Peer Community Journal* 2.

Fresneda J, Grebennikov VV, Ribera I. 2011. The phylogenetic and geographic limits of Leptodirini (Insecta: Coleoptera: Leiodidae: Cholevinae), with a description of Sciaphyes shestakovi sp. n. from the Russian Far East. *Arthropod Syst. Phylogeny*. 69.

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28:3150–3152.

Gibert J. Deharveng L. 2002. Subterranean ecosystems: a truncated functional biodiversity.

BioScience 52:473–481.

van Giesen L, Garrity PA. 2017. More than meets the IR: the expanding roles of variant Ionotropic Glutamate Receptors in sensing odor, taste, temperature and moisture. *F1000Res.* 6:1753.

Gomez-Diaz C, Bargeton B, Abuin L, Bukar N, Reina JH, Bartoi T, Graf M, Ong H, Ulbrich MH, Masson J-F, et al. 2016. A CD36 ectodomain mediates insect pheromone detection via a putative tunnelling mechanism. *Nature Communications* 7.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* 35:518–522.

Hüppop K. 2012. Adaptation to low food. In: White WB, Culver DC, editors. Encyclopedia of caves. Academic Press, Waltham, MA, p. 1–9.

Joseph RM, Carlson JR. 2015. Drosophila Chemoreceptors: A Molecular Interface Between the Chemical World and the Brain. *Trends Genet.* 31:683–695.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.

Knecht ZA, Silbering AF, Cruz J, Yang L, Croset V, Benton R, Garrity PA. 2017. Ionotropic Receptor-dependent moist and dry cells control hygrosensation in. *Elife* 6.

Kováč Ľ. 2018. Caves as Oligotrophic Ecosystems. In: Moldovan OT, Kováč Ľ, Halse S, editors. Cave Ecology. Cham: Springer International Publishing. p. 297–307.

Kozma MT, Ngo-Vu H, Wong YY, Shukla NS, Pawar SD, Senatore A, Schmidt M, Derby CD. 2020. Comparison of transcriptomes from two chemosensory organs in four decapod crustaceans reveals hundreds of candidate chemoreceptor proteins. *PLoS One* 15:e0230266.

Langille BL, Hyde J, Saint KM, Bradford TM, Stringer DN, Tierney SM, Humphreys WF, Austin AD, Cooper SJB. 2021. Evidence for speciation underground in diving beetles (Dytiscidae) from a subterranean archipelago. *Evolution* 75:166–175.

Lauritzen S-E. 2018. Physiography of the Caves. In: Moldovan OT, Kováč Ľ, Halse S, editors. Cave Ecology. Cham: Springer International Publishing. p. 7–21.

Lawrence JF. 2016. 10.4. Scirtidae Fleming, 1821. In: Beutel RG, Leschen RAB (eds) Handbook of zoology, vol 4/38: Coleoptera, beetles, vol 1. Morphology and systematics (Archostemata, Adephaga, Myxophaga, Polyphaga partim), 2nd edn. W DeGruyter, Berlin, pp 215–225

Leijs R, van Nes EH, Watts CH, Cooper SJB, Humphreys WF, Hogendoorn K. 2012. Evolution of blind beetles in isolated aquifers: a test of alternative modes of speciation. *PLoS One* 7:e34260.

Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49:W293–W296.

Librado P, Vieira FG, Rozas J. 2012. BadiRate: estimating family turnover rates by likelihood-based methods. *Bioinformatics* 28:279–281.

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol. Biol. Evol.* 38:4647–4654.

Mayer ML, Armstrong N. 2004. Structure and Function of Glutamate Receptor Ion Channels.

Annual Review of Physiology 66:161–181.

McKenna DD, Shin S, Ahrens D, Balke M, Beza-Beza C, Clarke DJ, Donath A, Escalona HE, Friedrich F, Letsch H, et al. 2019. The evolution and genomic basis of beetle diversity. *Proc. Natl. Acad. Sci. U. S. A.* 116:24729–24737.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* 37:1530–1534.

Moldovan OT, Jalžić B, Erichsen E. 2004. Adaptation of the mouthparts in some subterranean Cholevinae (Coleoptera, Leiodidae). *Natura Croatica* 13:1–18.

Ngoc PCT, Greenhalgh R, Dermauw W, Rombauts S, Bajda S, Zhurov V, Grbić M, Van de Peer Y, Van Leeuwen T, Rouzé P, et al. 2016. Complex Evolutionary Dynamics of Massively Expanded Chemosensory Receptor Families in an Extreme Generalist Chelicerate Herbivore. *Genome Biol. Evol.* 8:3323–3339.

Nichols Z, Vogt RG. 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 Biochem. Mol. Biol.* 38:398–415.

Ni L. 2020. The Structure and Function of Ionotropic Receptors in Drosophila. *Front. Mol. Neurosci.* 13:638839.

Pelosi P, Iovinella I, Felicioli A, Dani FR. 2014. Soluble proteins of chemical communication: an overview across arthropods. *Front. Physiol.* 5:320.

Pelosi P, Iovinella I, Zhu J, Wang G, Dani FR. 2018. Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. *Biological Reviews* 93:184–200.

Poulson TL. 1963. Cave Adaptation in Amblyopsid Fishes. *American Midland Naturalist* 70:257.

Price MN, Dehal PS, Arkin AP. 2010. FastTree 2-approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490.

Ribera I, Fresneda J, Bucur R, Izquierdo A, Vogler AP, Salgado JM, Cieslak A. 2010. Ancient origin of a Western Mediterranean radiation of subterranean beetles. *BMC Evol. Biol.* 10:29.

Robertson HM, Baits RL, Walden KKO, Wada-Katsumata A, Schal C. 2018. Enormous expansion of the chemosensory gene repertoire in the omnivorous German cockroach Blattella germanica. *J. Exp. Zool. B Mol. Dev. Evol.* 330:265–278.

Robertson HM, Warr CG, Carlson JR. 2003. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 100:14537–14542.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press.

Sánchez-Gracia A, Vieira FG, Rozas J. 2009. Molecular evolution of the major chemosensory gene families in insects. *Heredity* 103:208–216.

Schoville SD, Chen YH, Andersson MN, Benoit JB, Bhandari A, Bowsher JH, Brevik K, Cappelle K, Chen M-JM, Childers AK, et al. 2018. A model species for agricultural pest genomics: the genome of the Colorado potato beetle, Leptinotarsa decemlineata (Coleoptera: Chrysomelidae). *Sci. Rep.* 8:1931.

Stegner MEJ, Stemme T, Iliffe TM, Richter S, Wirkner CS. 2015. The brain in three crustaceans

from cavernous darkness. *BMC Neurosci.* 16:19.

Stocker RF. 1994. The organization of the chemosensory system in Drosophila melanogaster: a review. *Cell Tissue Res.* 275:3–26.

Thoma M, Missbach C, Jordan MD, Grosse-Wilde E, Newcomb RD, Hansson BS. 2019. Transcriptome Surveys in Silverfish Suggest a Multistep Origin of the Insect Odorant Receptor Gene Family. Front. Ecol. Evol. 7

Varatharasan N, Croll RP, Franz-Odendaal T. 2009. Taste bud development and patterning in sighted and blind morphs of Astyanax mexicanus. *Dev. Dyn.* 238:3056–3064.

Vizueta J, Frías-López C, Macías-Hernández N, Arnedo MA, Sánchez-Gracia A, Rozas J. 2017. Evolution of Chemosensory Gene Families in Arthropods: Insight from the First Inclusive Comparative Transcriptome Analysis across Spider Appendages. *Genome Biol. Evol.* 9:178–196.

Vizueta J, Sánchez-Gracia A, Rozas J. 2020. BITACORA: A comprehensive tool for the identification and annotation of gene families in genome assemblies. *Mol. Ecol. Resour.* 20:1445–1452.

Wang H-C, Minh BQ, Susko E, Roger AJ. 2018. Modeling Site Heterogeneity with Posterior Mean Site Frequency Profiles Accelerates Accurate Phylogenomic Estimation. *Systematic Biology* 67:216–235.

Watts C.H.S & Leijs R. 2008. Review of the epigean species of Paroster SHARP, 1882, with descriptions of three new species, and phylogeny based on DNA sequence data of two mitochondrial genes (Coleoptera: Dytiscidae: Hydroporinae). *Koleopterol. Rundsch.* 78 :9–36.

Yang J, Chen X, Bai J, Fang D, Qiu Y, Jiang W, Yuan H, Bian C, Lu J, He S, et al. 2016. The Sinocyclocheilus cavefish genome provides insights into cave adaptation. *BMC Biol.* 14:1.

Yoshizawa M, Jeffery WR, van Netten SM, McHenry MJ. 2014. The sensitivity of lateral line receptors and their role in the behavior of Mexican blind cavefish (Astyanax mexicanus). *J. Exp. Biol.* 217:886–895.

Zhao Y-J, Li G-C, Zhu J-Y, 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.