

OPEN ACCESS

EDITED BY

Ana Depetris Chauvin,
Max Planck Institute for Chemical
Ecology, Germany

REVIEWED BY

George F. Obiero,
Technical University of Kenya, Kenya
Juan Pedro Wulff,
Texas A&M University College Station,
United States

*CORRESPONDENCE

Juan Hurtado
hurtado.juanp@gmail.com
Esteban Hasson
ehasson@ege.fcen.uba.ar

†These authors have contributed
equally to this work

SPECIALTY SECTION

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

RECEIVED 30 May 2022

ACCEPTED 19 July 2022

PUBLISHED 09 August 2022

CITATION

Rondón JJ, Moreyra NN, Pisarenco VA,
Rozas J, Hurtado J and Hasson E
(2022) Evolution of
the odorant-binding protein gene
family in *Drosophila*.
Front. Ecol. Evol. 10:957247.
doi: 10.3389/fevo.2022.957247

COPYRIGHT

© 2022 Rondón, Moreyra, Pisarenco,
Rozas, Hurtado and Hasson. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Evolution of the odorant-binding protein gene family in *Drosophila*

Johnma J. Rondón^{1,2}, Nicolás Nahuel Moreyra^{1,2},
Vadim A. Pisarenco^{3,4}, Julio Rozas^{3,4}, Juan Hurtado^{1,2*†} and
Esteban Hasson^{1,2*†}

¹Departamento de Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina, ²Instituto de Ecología, Genética y Evolución de Buenos Aires, Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires, Argentina, ³Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona (UB), Barcelona, Spain, ⁴Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona (UB), Barcelona, Spain

Odorant-binding proteins (OBPs) are encoded by a gene family involved in the perception of olfactory signals in insects. This chemosensory gene family has been advocated as a candidate to mediate host preference and host shifts in insects, although it also participates in other physiological processes. Remarkable differences in the OBP gene repertoire have been described across insect groups, suggesting an accelerated gene turnover rate. The genus *Drosophila*, is a valuable resource for ecological genomics studies since it comprises groups of ecologically diverse species and there are genome data for many of them. Here, we investigate the molecular evolution of this chemosensory gene family across 19 *Drosophila* genomes, including the *melanogaster* and *repleta* species groups, which are mostly associated with rotting fruit and cacti, respectively. We also compared the OBP repertoire among the closely related species of the *repleta* group, associated with different subfamilies of Cactaceae that represent disparate chemical challenges for the flies. We found that the gene family size varies widely between species, ranging from 39 to 54 candidate OBPs. Indeed, more than 54% of these genes are organized in clusters and located on chromosomes X, 2, and 5, with a distribution conserved throughout the genus. The family sizes in the *repleta* group and *D. virilis* (*virilis-repleta* radiation) were smaller than in the *melanogaster* group. We tested alternative evolutionary models for OBP family size and turnover rates based on different ecological scenarios. We found heterogeneous gene turnover rates (GR) in comparisons involving columnar cactus specialists, prickly pear specialists, and fruit dwellers lineages, and signals of rapid molecular evolution compatible with positive selection in specific OBP genes. Taking ours and previous results together, we propose that this chemosensory gene family is involved in host adaptation and hypothesize that the adoption of the cactophilic lifestyle in the *repleta* group accelerated the evolution of members of the family.

KEYWORDS

OBPs, chemosensory, cactophilic, comparative genomics, adaptation, *Drosophila*

Introduction

The radiation of angiosperms, which occurred ~130 My ago, is considered one of the main events underlying the evolutionary success and diversification of Arthropods (Grimaldi and Engel, 2005). Many insect families have evolved in association with plants, and these insect-plant interactions favored scenarios of ecological specialization in resource use (Forister et al., 2012). Insects are a group with amazing adaptability to different host plants, especially considering the remarkable ability of plants to produce and release a wide range of compounds with toxic effects (War et al., 2018).

In animals, environmental sensing is modulated by sensory perception systems that have been crucial in the conquest of new niches (Arguello and Benton, 2017). Sensory perception is a key step in animal feeding, reproduction, and survival. Many signals involve chemical molecules that are recognized with a complex and sophisticated chemosensory machinery that comprises olfaction and taste. In *Drosophila*, as well as in other insects, the antennae and maxillary palps are the main organs participating in the perception of chemical stimuli (Vogt and Riddiford, 1981; Pelosi, 1994). These sensory organs express proteins that detect odorant molecules and trigger the activation of molecular signaling pathways that may finally affect behavior (Hildebrand and Shepherd, 1997; Krieger and Ross, 2002). This complex system is found in vertebrates and invertebrates; however, their evolutionary origin differs between these distant groups, and many of the involved proteins are in fact non-homologous (Kaupp, 2010). In insects, protein families that are part of the chemosensory system are classified into two main groups: (A) soluble binding proteins like OBP, chemosensory proteins (CSP), Niemann-Pick type C2 (NPC2), and odorant-degrading enzymes (ODE); and (B) membrane-bound receptors, i.e., proteins that play a role as receptors in the dendrites of olfactory neurons. OBPs and CSPs participate in the recognition and transport of hydrophobic odorant molecules through the sensillar lymph, whereas receptors are responsible for the activation of intracellular signaling pathways in olfactory receptor neurons. Further, OBPs have been also detected in reproductive organs, where they participate in the recognition of molecules relevant to reproductive behavior (Shorter et al., 2016). The second group includes protein families such as odorant receptors (OR), gustatory receptors (GR), and ionotropic receptors (IR) (Pelosi, 1996; Joseph and Carlson, 2015).

Evolutionary studies of chemosensory families suggest that GR and IR emerged during the diversification of Metazoa, while the CSP family appeared later in arthropods, and the OR and OBP families originated more recently in insects (Eyun et al., 2017; Vizueta et al., 2020b). However, distant OBP-like homologs have been described in Araneae (Vizueta et al., 2017). The accelerated diversification reported for the OBP

family among arthropods triggered the interest in sensory-active genes and promoted comparisons with other animals. For instance, proteins with functions similar to OBPs reported in vertebrates received the same family name, although they are not evolutionarily related (Robertson, 2019).

Insect chemosensory gene families exhibit rapid evolutionary rates (Sánchez-Gracia et al., 2009; Brand et al., 2015; Diaz et al., 2018; Robertson, 2019). The size of the OBP family, i.e., the number of copies, varies widely across insect groups, including *Drosophila* (Vieira and Rozas, 2011). Expansion and contraction of gene families are the result of changes in the gene repertoire that may be caused by several mechanisms, e.g., paralog duplication, deletion, or pseudogenization. Such dynamics of gains and losses within a gene family are thought to be influenced by ecological factors, including the use of alternative hosts and resources. Thus, ecological changes, such as host shifts, may alter gene turnover rates (GR) in chemosensory gene families (e.g., Almeida et al., 2014 and references therein). Ecological adaptation is not only expected to affect GR but may also trigger adaptive divergence between ortholog CSP. Indeed, there is evidence of positive selection shaping the evolution of chemosensory genes between recently diverged ecologically different species (Kulmuni et al., 2013; Diaz et al., 2018).

Currently, more than 100 high-quality *Drosophila* spp. genomes have been published, and this number is constantly increasing, offering unprecedented opportunities for comparative genomics and evolutionary studies (Kim et al., 2021). However, the ecology of most species is poorly known and in many cases, differs greatly from *D. melanogaster*, the standard reference in genomic studies (Goldman-Huertas et al., 2015). Thus, we need to be cautious when extrapolating knowledge from *D. melanogaster* to other species.

One species group with a well-studied ecology, that markedly differs from species of the *melanogaster* group, is the *repleta* group. It includes more than 100 species, most adapted to a cactophilic lifestyle, i.e., they use necrotic cacti as breeding, mating, and feeding sites (Oliveira et al., 2012; Jezovit et al., 2017; O'Grady and DeSalle, 2018a). Within this group, the genomes of several species have been sequenced: *D. hydei* (Rane et al., 2019), *D. mojavensis*, *D. arizona*, and *D. navojoa* (*mojavensis* cluster, *mulleri* complex) (Clark et al., 2007; Sanchez-Flores et al., 2016), and *D. buzzatii*, *D. boroborema*, *D. koepferae*, and *D. antonietae* (*buzzatii* cluster, *mulleri* complex) (Guillén et al., 2014; Rane et al., 2019; Moreyra et al., 2022). Ecological specialization and resource use have been advocated as key aspects in the evolution and diversification of both the *mojavensis* and the *buzzatii* clusters (Matzkin, 2014; Etges, 2019; Hasson et al., 2019; Markow, 2019). For instance, *D. navojoa* and *D. buzzatii* are specialized in the use of prickly pears of the genus *Opuntia* (subfamily Opuntioideae), the ancestral state of host plant use, whereas other species mainly

use columnar cacti of the subfamily Cactoideae (Oliveira et al., 2012; Hasson et al., 2019).

In this study, we analyzed *Drosophila* species of the *melanogaster* and the *repleta* groups as representatives of lineages using decaying fruits and cacti, respectively, as breeding sites. We aim to investigate the evolutionary changes in the OBP gene family likely associated with the acquisition of cactophily in the *repleta* group and the more recent transition from prickly pears to columnar cacti in specific clades of the *repleta* group. Our main hypothesis is that the repertoire of OBP genes changed in concert with the adoption of different lifestyles and host use.

Materials and methods

Identification of odorant-binding protein gene candidates in *Drosophila* genomes

We used genomic data of 18 *Drosophila* species (19 genomes, including two *D. koepferae* genomes) as input for BITACORA v.1.2.1 software (Vizueta et al., 2020c) to search for gene candidates encoding OBPs. To this end, we included several species of the *melanogaster* group (*D. simulans*, *D. sechellia*, *D. erecta*, *D. yakuba*, *D. suzukii*, and *D. melanogaster*) and *repleta* group (*D. buzzatii*, *D. koepferae*, *D. borborema*, *D. antonietae*, *D. navojoa*, *D. arizonae*, *D. mojaviensis*, and *D. hydei*) with well-annotated genomes. Representatives of other groups, i.e., *D. pseudoobscura*, *D. virilis*, *D. willistoni*, and *D. grimshawi*, were also included to obtain information outside of the focus groups (Table 1 and Figure 1). For each analyzed species, BITACORA requires the genome, its annotation in General Feature Format (GFF file), and the reference proteome. All these files were obtained from the National Center for Biotechnology Information (NCBI) (Sayers et al., 2021) and FlyBase (Larkin et al., 2021), except for species of the *buzzatii* cluster consisting of two strains of *D. koepferae* (Fontdevila and Wasserman) derived from collections in Argentina (A) and Bolivia (B), *D. buzzatii* (Patterson and Wheeler), *D. antonietae* (Tidon-Sklorz and Sene), and *D. borborema* (Vilela and Sene); which were recently reported in Moreyra et al. (2022) (Supplementary Table 1). We used the BITACORA pipeline to perform two rounds of similarity-based searches with BLAST v. 2.9.0-2 (Camacho et al., 2009) and HMMER v3.3 (Eddy, 2011) as recommended in Vizueta et al. (2020a). In the first round, databases and HMM profiles created with HMMER of known OBPs in insects (i.e., *D. melanogaster*, *Apis mellifera*, and *Tribolium castaneum*) were employed to detect candidate genes (Vizueta et al., 2020a). We set a relaxed *E*-value threshold of 1×10^{-3} for these searches since our databases included OBP amino acid sequences of lineages highly divergent from Drosophilidae. In the second round, we used a new

database containing exclusively the *Drosophila* candidate OBP proteins identified in the first round and their respective annotations. This last round allowed more exhaustive searches that could lead to the discovery of new genes. The new OBP candidates were filtered with a minimum *E*-value of 1×10^{-8} . We applied this pipeline for each species to obtain the final set of species-specific candidate sequences of the OBP family. To improve the annotation of OBP gene candidates in cactophilic species of the *repleta* group, we used transcriptomic data of the adult whole body available in the European Nucleotide Archive (ENA) at EMBL-EBI under accession numbers PRJNA540063,¹ PRJNA414017,² and PRJNA395148.³ To edit and visualize the genomic information we used the Apollo genome browser v.2.0.8 (Lee et al., 2009). The manual curation of candidate gene models allowed to reduce biases produced by annotation heterogeneity (Weisman et al., 2022). Also, to validate each OBP candidate, we identified OBP functional domains with Interproscan v.5.52 (Jones et al., 2014) by comparing candidate OBPs against the Pfam database v.33.1 (Mistry et al., 2021) and Conserved Domains Database (CDD) v.3.18 (Marchler-Bauer et al., 2005). Lastly, we counted the number of disulfide bridges, conserved cysteine residues, and alpha-helix motifs in each candidate protein.

Organization and localization of odorant-binding protein genes in the *repleta* group

To describe the genetic organization of candidate OBPs in each species of the *repleta* group, we computed nucleotide sequence length and number of exons. In *D. melanogaster* the clustering of OBP genes is already known (Librado and Rozas, 2013), thus, we evaluated the degree of gene clustering in other species by analyzing the physical distance among OBPs located in the same contig/scaffold. To define the clusters we followed the criteria proposed by Vieira et al. (2007). Thus, we consider that a given number of OBP genes are clustered if they are in a region of length $C_L = g(n-1)$, considering a value of $g = 10$ kb (based on gene density of *D. melanogaster*). To map clusters, we performed local BLASTN searches using the chromosome-level assembly of *D. buzzatii* as subject (Guillén et al., 2014) and the scaffolds of each species of the *repleta* group as queries. Hits with the best score and *E*-value allowed to identify the scaffold-chromosome correspondence. The location of clusters in species of the *repleta* group was compared with that reported in *D. melanogaster*.

1 <https://www.ebi.ac.uk/ena/browser/view/PRJNA540063>

2 <https://www.ebi.ac.uk/ena/browser/view/PRJNA414017>

3 <https://www.ebi.ac.uk/ena/browser/view/PRJNA395148>

TABLE 1 Numbers of OBP genes identified in the *Drosophila* species included in this study.

Species*	Subgenus	Species group	Total unique OBPs** (1st + 2nd round)
<i>D. antonietae</i> (Dato)	<i>Drosophila</i>	<i>repleta</i>	42
<i>D. borborema</i> (Dbrb)	<i>Drosophila</i>	<i>repleta</i>	48
<i>D. buzzatii</i> (Dbuz)	<i>Drosophila</i>	<i>repleta</i>	44
<i>D. koepferae A</i> (DkoeA)	<i>Drosophila</i>	<i>repleta</i>	44
<i>D. koepferae B</i> (DkoeB)	<i>Drosophila</i>	<i>repleta</i>	42
<i>D. mojavensis</i> (Dmoj)	<i>Drosophila</i>	<i>repleta</i>	42
<i>D. navojoa</i> (Dnav)	<i>Drosophila</i>	<i>repleta</i>	43
<i>D. arizonae</i> (Dari)	<i>Drosophila</i>	<i>repleta</i>	39
<i>D. hydei</i> (Dhyd)	<i>Drosophila</i>	<i>repleta</i>	49
<i>D. virilis</i> (Dvir)	<i>Drosophila</i>	<i>virilis</i>	40
<i>D. grimshawi</i> (Dgri)	<i>Drosophila</i>	<i>grimshawi</i>	50
<i>D. melanogaster</i> (Dmel)	<i>Sophophora</i>	<i>melanogaster</i>	52
<i>D. sechellia</i> (Dsec)	<i>Sophophora</i>	<i>melanogaster</i>	52
<i>D. simulans</i> (Dsim)	<i>Sophophora</i>	<i>melanogaster</i>	52
<i>D. erecta</i> (Dere)	<i>Sophophora</i>	<i>melanogaster</i>	50
<i>D. yakuba</i> (Dyak)	<i>Sophophora</i>	<i>melanogaster</i>	53
<i>D. sukukii</i> (Dsuz)	<i>Sophophora</i>	<i>melanogaster</i>	53
<i>D. pseudoobscura</i> (Dpse)	<i>Sophophora</i>	<i>obscura</i>	44
<i>D. willistoni</i> (Dwil)	<i>Sophophora</i>	<i>willistoni</i>	54

*FlyBase abbreviations for *Drosophila* species names are shown in parenthesis. **The total of unique genes includes OBPs detected in the two rounds.

Sequence comparison of odorant-binding proteins across *Drosophila* species

The identification of OBP gene orthologs across 19 genomes was conducted using OrthoFinder v.2.5.2 (Emms and Kelly, 2019). Then, based on these searches, we calculated the number of orthologous sequences within each species group, i.e., *melanogaster*, *repleta*, and *virilis*. Lastly, using the species of the *virilis-repleta* radiation and *D. melanogaster*, we performed large-scale comparative genomic analyses with compareM⁴ to estimate the average amino acid identity (AAI) in both all annotated proteins (excluding OBPs) and only OBPs.

Gene turnover in the odorant-binding protein gene family

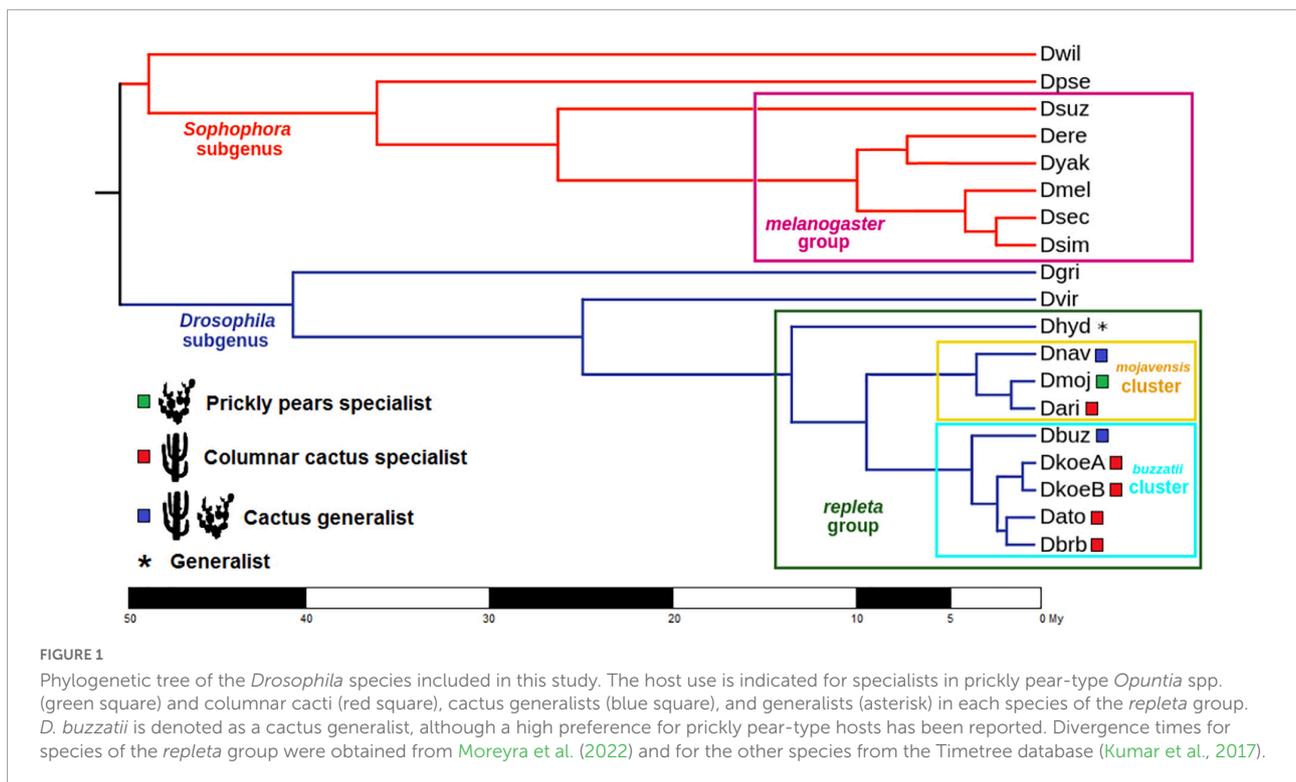
We studied gene family dynamics by means of GR. We first constructed an ultrametric species tree based on 3,405 single-copy orthologs (1:1 orthogroups) obtained from the OrthoFinder output. To this end, orthogroups were aligned

with MAFFT v.7.453 (Katoh and Standley, 2013) and misaligned regions were trimmed with TrimAl v.14 (Capella-Gutiérrez et al., 2009). Then, we concatenated the alignments with FASconCAT-G v.1.05 (Kück and Meusemann, 2010) and estimated branch lengths using RAXML v.8.2.12 (Stamatakis, 2014) under a partition model. For each orthogroup, we estimated the best substitution model with PROTTEST v.3.4.2 (Darrriba et al., 2011). The tree was calibrated with the R8S program v.1.81 (Sanderson, 2003) by setting a time of 50 million years to the ancestral node (root), the split between the *Sophophora* and *Drosophila* subgenera, obtained from the Timetree database (Kumar et al., 2017). Then, we estimated the turnover rates in the OBP family by means of maximum likelihood (ML) searches implemented in BadiRate v.1.35 (Librado et al., 2012). This program allows the estimation of turnover rates such as birth (β), gain (γ), death (δ), and innovation (i). Birth (e.g., a gene gain by unequal crossing-over) and death (e.g., a gene loss via deletion or pseudogenization) rates are considered as density-dependent, i.e., the probabilities of gene birth and death are proportional to the actual number of genes. On the other hand, gain and innovation rates [e.g., horizontal gene transfer (HGT) or *de novo* origin of short cis-regulatory elements] are considered independent of the actual family size. We used different stochastic population models, which we denote as model classes. Briefly, the birth-death (BD) and gain-death (GD) population model classes assume different rates for gene birth/gain and death events. Specifically, the BD estimates gene birth and death events, while the GD estimates gene gains and death events. The birth-death-innovation (BDI) population model class, in turn, admits three different rates, adding to the BD model class a rate for gene innovation events. Finally, the model class lambda-innovation (LI) represents a particular scenario of the BDI model class, assuming a unique rate for gene birth and death events.

We started with an approach that estimated global turnover rates for each model class described above assuming that all branches of the phylogenetic tree have the same rate(s). This set of models was termed *Global Rates* (GR) models. We compared the statistical fit between model classes using the Likelihood Ratio Test (LTR) and the Akaike Information Criterion (AIC). Since the BDI model class was the one showing the best fit to the data (see “Results” section), it was chosen for the estimation of turnover rates in the next approach in which rate heterogeneity was allowed for specific clades. This second set of models was termed *Branch Rates* (BR) models. Then, we used the *-bmodel* option (BDI-BR-ML approach) to run five BR models. Thus, groups of specific clades admitting different turnover rates were selected for each BR model according to the host range reported for the involved species or clades (Figure 1).

The first model (*Mmel*) admitted different rates between the branches of the *melanogaster* group and the remaining branches in the tree. The second model (*Mrep*) allowed different rates between the branches of cactophilic *repleta* group

⁴ <https://github.com/dparks1134/CompareM>



and the remaining branches. The third model (*Mmel-rep*) admitted different rates for the branches of the *melanogaster* group, the branches of the cactophilic *repleta* group, and the rest of the tree branches. The fourth model (*Mspe*) accepted different rates for the *repleta* branches involving lineages of columnar-cactus specialists, the remaining *repleta* branches excluding the generalist *D. hydei*, and the rest of the branches. The fifth model admitted different turnover rates for the *melanogaster* group branches, the *repleta* lineages of columnar-cactus specialists, *repleta* branches of prickly pear specialists, and the rest of the branches. We used the BDI-GR model obtained with the global-rates as a null model (*Mnul*). We schematize all BR models tested in **Supplementary Figure 1**.

In addition, we searched for candidate OBP pseudogenes in the *repleta* group. First, we selected orthogroups (created with OrthoFinder) that have more copies of *D. melanogaster* than of *repleta* group species, which is expected for OBPs that went through pseudogenization in the *repleta* lineage. Then, for each selected OBP (*D. melanogaster* OBP belonging to any of the selected orthogroups), we identified the flanking conserved genes as “flags” and aligned them to the cactophilic genomes. This procedure allowed us to identify the syntenic region of each selected OBP in each *repleta* species genome. Lastly, we performed TBLASTN searches of each selected OBP against the respective syntenic regions identified in the *repleta* genomes and manually look into the obtained alignments to detect traces of pseudogenization events. We considered multiple hits

(alignment bit score > 20) on partial ORFs, having the same orientation and covering different portions of the OBP used as query, as indicative of the presence of a pseudogene.

Odorant-binding protein gene evolution in the *repleta* group

To explore kinship relationships between OBPs, we constructed ML phylogenetic trees with IQTree v.1.6.12 (Nguyen et al., 2015) using amino acid sequences of the seven cactophilic species of the *repleta* group and *D. melanogaster* as outgroup. Considering that our approach pointed to the interspecific level, we decided to exclude the sequences of *D. koepferae* B from this analysis. Sequences were first aligned with MAFFT and the presence of conserved cysteine residues, typical of OBPs, was confirmed by manual inspection of each alignment. The OBP trees were customized using iTOL v.6.5.2 (Letunic and Bork, 2021) and plotted showing functional domains including only three species: *D. borborema*, *D. mojavensis* (as representatives of the respective clusters), and *D. melanogaster*. OBPs were classified into subfamilies based on a phylogenetic criterion. Thus, any sequence included in the clades previously identified in Vieira et al. (2007) was assigned to the corresponding subfamily. Then, the classification was validated in each clade by comparing the number of conserved cysteines and functional domains commonly found in each subfamily.

We employed DnaSP v.6.12.03 (Librado and Rozas, 2009) to estimate nucleotide sequence divergence within the *repleta* group, the *buzzatii* cluster, and the *mojavensis* cluster, using *D. virilis*, *D. mojavensis*, and *D. buzzatii* as outgroups in each comparison, respectively. Particularly, we estimated mean nucleotide divergence (K) using Jukes and Cantor (JC) distance, and the ratio of non-synonymous to synonymous substitution rates ($\omega = Ka/Ks$). Then, we performed gene-wide and codon-based selection tests in orthogroups containing at least one OBP per species (seven species of the *repleta* group plus *D. melanogaster* and *D. virilis*). Therefore, we used different methods available in the HyPhy v.2.5.31 package (Kosakovsky Pond et al., 2020). These methods consists of searches of selection footprints on branches (aBSREL) and sites (MEME), and a non-site-specific method (BUSTED) that reports evidence of gene-wide episodic diversifying selection on at least one site from at least one branch of the orthogroup phylogeny (Murrell et al., 2012, 2015; Smith et al., 2015). Moreover, we tested for the fit of codon models M1a, M2a, M7, and M8 with Codeml available in PAML v.4.9 package (Yang, 2007). For these analyses, we aligned coding sequences of each orthogroup with TranslatorX software (Abascal et al., 2010) and built ML trees with RAxML v.8.2.12.

Results

Odorant-binding protein gene repertoire varies among *Drosophila* lineages

The number of candidate OBPs identified ranged from 39 to 54 for the species included in this study (Table 1). Differences in family size were found between species within both the *melanogaster* group (range = 50–53 genes) and the *virilis-repleta* radiation (range = 39–49 genes) (Table 1). Additionally, it was striking to find a slight difference in the number of OBPs at the intra-specific level, between *D. koepferae* strains A and B.

Odorant-binding proteins are organized in clusters

We identified 7–10 OBP gene clusters per species in the *repleta* group, each one harboring 2–7 genes (Table 2). We tested the relationship of the number or length of scaffolds with the number of OBP gene clusters to evaluate whether assembly contiguity influences the number of clusters that can be detected in each species. Correlation analyses yielded non-significant results for all *repleta* species (Pearson's correlation P -values ranged from 0.389 to 0.504). Moreover, more than half (54.76–66.67%) of the OBP genes were found in clusters in all species, which indicates a non-random distribution of

OBP genes throughout the genome (Table 2). We found that most OBP genes are located on Muller elements A, C, and E in the *repleta* group (see Table 2). The Muller C element, which represents chromosome 5 and arm 2R in the *repleta* group and in *D. melanogaster*, respectively, harbor most OBPs and clusters in all species. This chromosome shows about four times more genes and clusters than, for instance, the X chromosome (Muller element A). We found that the localization, organization, and subfamily content of OBP gene clusters are conserved between the *repleta* group and *D. melanogaster*. The most conserved *D. melanogaster* gene clusters were those of Obp19, Obp50, Obp56, Obp57, Obp58, Obp83, and Obp99 (Table 2 and Figure 2).

Comparison of odorant-binding protein sequences among the *repleta* group

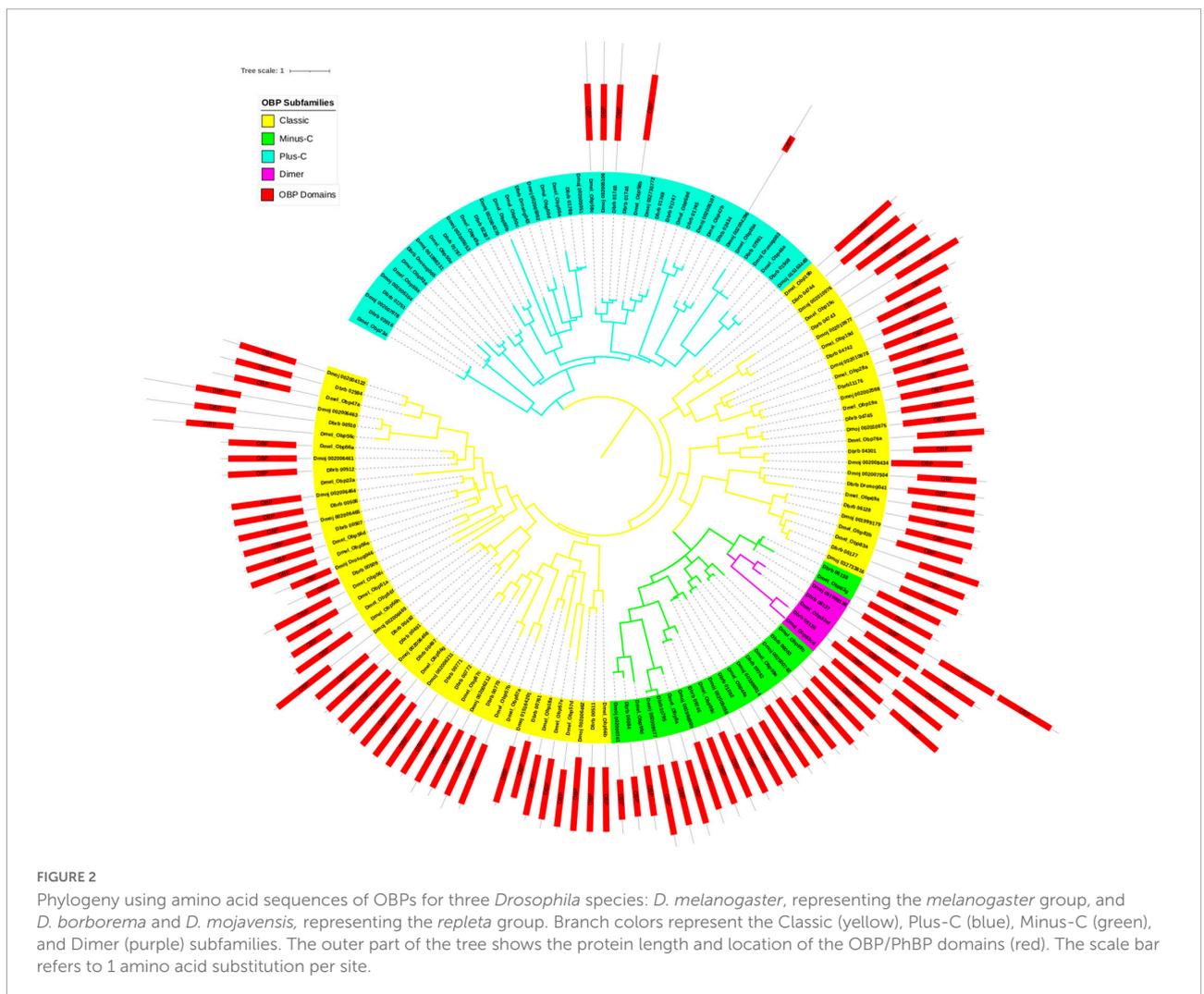
We also compared gene and protein features of OBPs across species of the *repleta* group. First, we detected from one to eight exons per gene (Supplementary Figure 2), and second, clustering analyses showed, as expected, that more divergent species share fewer orthologs than more closely related ones. In fact, 35 OBP orthogroups were shared among all descendants of the genus *Drosophila* (tree root), and 39 among branches of the *repleta* group (Supplementary Figure 3). We also compared sequence divergence in the OBP repertoire and the entire proteome (excluding OBPs) computing average pairwise AAI between species. For both the proteome and OBP repertoire, sequence divergence was higher in phylogenetically distant species than in more closely related ones (Supplementary Tables 2, 3). Also, we found greater divergence in OBP sequences than in the remaining proteome (T -test P -value 0.025) (Supplementary Figure 4).

Cactophilic and specialist lineages of the *repleta* group have distinctive gene turnover rates

We estimated global GR to compare among stochastic population model classes BDI, BD, GD, and LI. The best fit to the data was obtained with BDI (BDI-GR-ML combination) according to the results of LTR (P -value < 0.001) and AIC (Table 3). These analyses showed that the birth (β) rate was higher than the death (δ) rate, indicating an expansion of the OBP family during the radiation of the *Drosophila* genus. As expected, the innovation rate was close to zero and substantially lower than both the birth and death rates (Supplementary Table 4). Likewise, a minimum number of gains and losses were estimated using global and BR, throughout the *Drosophila* phylogeny (Figure 3).

TABLE 2 Number of clusters and OBP genes (in parenthesis) across the Muller's Elements (chromosomes) of the *repleta* group species.

Species	Muller's element (chromosome)						Total number of genes in clusters	Percentage of genes in clusters
	A (X)	E (2)	B (3)	D (4)	C (5)	F (6)		
<i>D. antonietae</i>	1 (4)	2 (4)	0	0	4 (15)	0	23	54.76
<i>D. borborema</i>	1 (4)	3 (8)	0	0	5 (20)	0	32	66.67
<i>D. buzzatii</i>	1 (4)	4 (9)	0	0	4 (16)	0	29	65.9
<i>D. koepferae A</i>	1 (4)	2 (4)	0	1 (3)	4 (18)	0	29	65.9
<i>D. koepferae B</i>	1 (4)	3 (6)	0	0	4 (16)	0	26	61.9
<i>D. mojavensis</i>	1 (4)	3 (6)	0	0	4 (16)	0	26	61.9
<i>D. arizonae</i>	1 (4)	3 (6)	0	0	3 (13)	0	23	58.97
<i>D. navojoa</i>	2 (8)	4 (8)	0	0	3 (11)	0	27	62.79
<i>D. melanogaster</i>	1 (4)	3 (8)	0	0	6 (21)	0	33	63.46



In the BDI-BR-ML approach, we investigated putative heterogeneous OBP turnover rates across specific lineages in the phylogeny to explore whether the type of host, particularly the

acquisition of cactophily, affected turnover rates. All evaluated models showed a better fit than the null model (*Mnul*), suggesting that GR have been influenced by the acquisition of

TABLE 3 Likelihood and AIC values for the birth-death-innovation (BDI), birth-death (BD), lambda-innovation (LI), and gain-death (GD) models assuming global (GR) or branch-specific rates (BR), and with maximum likelihood (ML) inference.

Method	Model	np	L	AIC
<i>Global rates</i>				
BDI-ML	GR	4	−368.82*	745.64
BD-ML	GR	3	−407.4	820.8
LI-ML	GR	2	−409.41	822.81
GD-ML	GR	4	−398.52	805.04
<i>Branch rates</i>				
BDI-BR-ML	<i>Mnul</i>	4	−368.82	745.64
	<i>Mmel</i>	7	−367.43	748.86
	<i>Mrep</i>	7	−360.1	734.2
	<i>Mmel-rep</i>	10	−359.56	739.12
	<i>Mspe</i>	10	−348.08**	716.16
	<i>Mall</i>	13	−347.51**	721.08

*Best GR model class (LRT, P -value < 0.001). **Best BR model (LRT, P -value < 0.001). np is number of parameters. Best values of the Akaike Information Criterion are in bold.

the cactophilic lifestyle (Table 3 and Supplementary Table 5). In particular, the *Mspe* and *Mall* models showed the highest goodness of fit (LRT, P -value < 0.001). These models involved different turnover rates for branches of fruit dwellers, columnar cacti specialists, and prickly pears specialists (Table 3 and Supplementary Table 5).

We also explored the possible role of pseudogenization as a mechanism responsible for gene loss in the OBP family. Five *D. melanogaster* OBPs, namely *obp56i*, *obp57b*, *obp57d*, *obp57e*, and *obp22a* could not be detected as annotated proteins in the genomes of the *repleta* group species. The searches of these OBPs in the respective syntenic regions identified in the *repleta* genomes revealed no detectable traces of pseudogenization except for *obp51a*. For this OBP, we detect pseudogenization traces in all the inspected cactophilic species except for *D. arizonae*. In *D. melanogaster*, the *obp51a* gene is located between the CG7639 and *chn* genes, a region of approximately 160 Kb (Supplementary Table 6).

Odorant-binding protein gene evolution in the *repleta* group

Four protein subfamilies can be distinguished in the OBP tree of the *repleta* group: Classic, Minus-C, Plus-C, and Dimer (Figure 2). The former has the typical OBP structure with six conserved cysteines, at least one proline, and 4–6 alpha helix motifs. The Plus-C subfamily has a greater number of cysteines and alpha-helix motifs than the Minus-C subfamily. The Dimer subfamily showed more than one functional OBP/Pheromone binding Protein (PhBP) domain. We found a striking level of divergence in the Plus-C subfamily, for which many sequences lacked the typical OBP hydrophobic ligand-binding domain.

Figure 4A shows the average nucleotide sequence divergence (K -JC) in pairwise comparisons between species of the *repleta* group and *D. virilis*, between species of the *buzzatii* cluster and *D. mojavensis*, and between species of the *mojavensis* cluster and *D. buzzatii* (Figure 4A). In general, ω -values were lower than 1 for most OBP gene, indicating negative selection as the main force shaping OBP evolution. However, we also detected OBPs showing clear signals of positive selection (Figures 4B,C). To further explore these cases, we employed methods implemented in the Hyphy and PAML packages to detect sites showing signals of positive selection in every branches. Overall, tests of adaptive evolution yielded significant results, pointing to several OBPs as targets of selection. Positive selection signals were found in 9 of 31 orthogroups in at least one branch of the *repleta* group (aBSREL method), particularly for the terminal branches of *D. borborema*, *D. buzzatii*, and *D. mojavensis*. The strongest signals of selection were detected for orthogroups *obp56a*, *obp44a*, *obp83a*, *obp8a*, and *obp84a*, which showed evidence of selection with at least two of the tested methods implemented in Hyphy (Supplementary Table 7).

Discussion

In this study, we investigate the evolution of the OBP family in the genus *Drosophila* by comparing the *melanogaster* group and the *virilis-repleta* radiation with emphasis on cactophilic species of the *repleta* group. The main aim of our study was to evaluate whether there is a relationship between the evolution of OBPs and the acquisition of cactophily in the *repleta* group. We identified candidate OBPs and estimated GR along branches of the species tree. This is the first study, as far as we know, reporting the organization and evolution of this family of chemosensory genes in species of the *repleta* group, other than *D. mojavensis*, in a phylogenomic framework.

The analyses of putative OBP genes identified in each genome revealed that the size of the OBP family in the subgenus *Sophophora* is close to the number reported in *D. melanogaster* (Vieira et al., 2007; Vieira and Rozas, 2011) and slightly greater than in the *virilis-repleta* radiation (subgenus *Drosophila*), in which the number of OBPs was only known in *D. grimshawi*, *D. virilis*, and *D. mojavensis*. Successive nested rounds of OBP searches improved the search precision of gene families in non-model species as recommended by Vizueta et al. (2020a). Following this approach, OBP repertoires identified were in close proximity to numbers previously reported in *Drosophila* (Vieira et al., 2007; Almeida et al., 2014).

Drosophila OBP genes have a simple organization, typically consisting of 1–4 exons distributed in clusters (more than 54%) throughout the genomes. Such clustering is commonly found in genes related by function and regulation patterns (Yi et al., 2007). The OBPs in species of the *repleta* group,

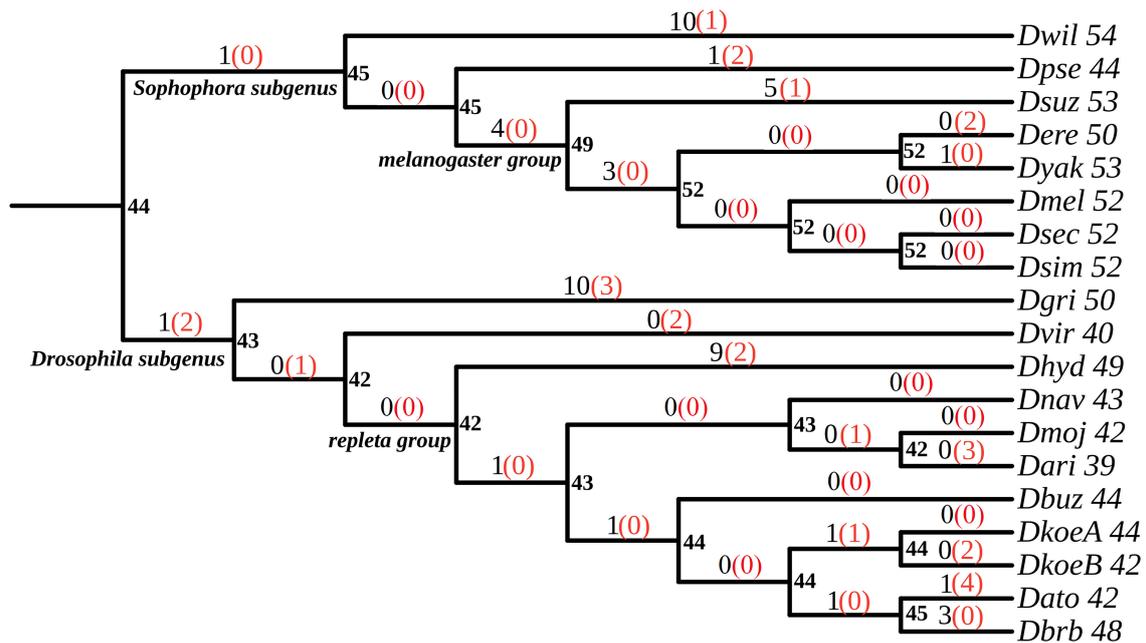


FIGURE 3 Phylogenetic tree showing OBP births (gains) and losses in *Drosophila* species. Values were obtained with BadiRate under the birth-death-innovation (BDI) model class with the BDI-GR-ML approach. Numbers next to the species names represent the OBP family size. Numbers at internal nodes represent the estimated family size of each ancestor. Numbers on each branch stand for the estimated gene gains (black) and losses (red, in parentheses).

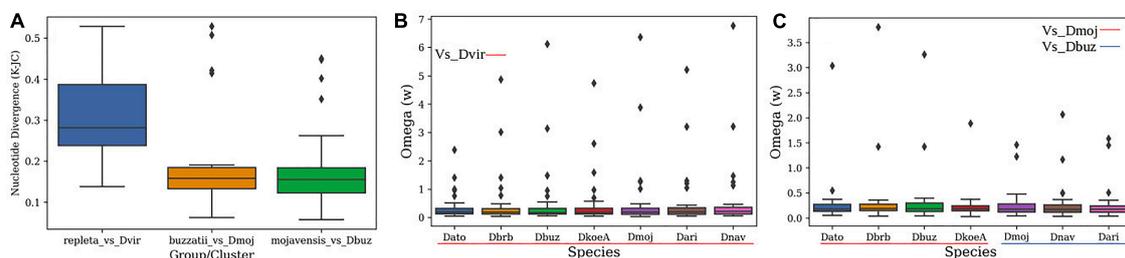


FIGURE 4 Distribution and patterns of nucleotide divergence (K-JC) for OBP genes. (A) Nucleotide divergence of the *repleta* group, *buzzatii* cluster, and *mojavensis* cluster. (B) ω -values in pairwise comparisons between *mulleri* subgroup species and *D. virilis*, (C) ω -values in pairwise comparison between species of the *buzzatii* cluster and *D. mojavensis*, and between species of the *mojavensis* cluster and *D. buzzatii*.

are mainly distributed on Muller elements A, E, and C as in *D. melanogaster*. These elements correspond to chromosomes X, 2, and 5 in the *repleta* group and to chromosome X and arms 3R and 2R in the *melanogaster* group. We also found a higher density of OBP genes and clusters on chromosome 5 than in the other elements, a feature conserved across all species analyzed. Even though rates of chromosomal rearrangements reported in *Drosophila* are the highest among eukaryotes, most genes remained in the same chromosome arms (Ranz et al., 1997; Richards et al., 2005; Schaeffer, 2018). However, the order and orientation of genes changed across species and lineages, likely due to chromosomal inversions, which have been key in

preventing gene flow and promoting speciation (Ranz et al., 2001). This evidence points to paracentric inversions (by far the most frequent rearrangements in the genus), as the dominant mechanism for shuffling the order of genes within chromosomes (Bhutkar et al., 2008). Therefore, our results give support to previous explanations proposed to account for the conserved localization of OBP genes and clusters across species of the subgenus *Sophophora*.

Protein sequences encoded by single-copy genes were employed to build a species tree that showed the same topology as in recent phylogenomic studies of the *repleta* group (Moreyra et al., 2022) and the *buzzatii* cluster (Hurtado et al., 2019). Our

estimation of GR using the BDI model, suggests an expansion of the OBP family in *Drosophila*. Furthermore, we found differences in patterns of gene gain and gene loss events between the *melanogaster* group (fruit dwellers) and cactophilic lineages of the *repleta* group. The former had an accelerated expansion in the last 50 My and accumulated between 7–9 lineage-specific net gene gains. This increase in family size appears to be coincident with the continued radiation of insects during the Cenozoic, that is probably linked to the success of flowering plants (Grimaldi and Agosti, 2000). In fact, studies of insect-plant interactions suggest that the diversification of angiosperms opened new niches, accelerating speciation rates in insects (van der Niet and Johnson, 2012; Zheng et al., 2018; Silvestro et al., 2021; Wang et al., 2022). On the other hand, our analyses indicated that the *repleta* group presumably experienced OBP gene losses that probably occurred during the acquisition of cacti, for instance via pseudogenization of genes involved in fruit recognition. Similar reductions in family size were reported for OR and GR receptors in two specialists of the *melanogaster* subgroup: *D. sechellia* and *D. erecta* (McBride, 2007; McBride and Arguello, 2007; Gardiner et al., 2008). However, we did not detect evidence of pseudogenization in the *repleta* group, except for *obp51a*, for which traces were detectable in almost all inspected genomes of cactophilic species. Pseudogenes detection may be hindered by the rapid divergence expected for non-coding sequences. Therefore, further efforts should be oriented to design strategies to investigate whether the absence of OBPs in the *repleta* group most likely reflects independent gains in the *Sophophora* subgenus branches or ancestral pseudogenization events in the *repleta* lineage.

The BDI model allowed us to consistently unveil the evolutionary dynamics of the OBP family, as reported for other chemosensory families like odorant (ORs), gustatory (GRs), and ionotropic (IRs) receptors (Hahn et al., 2005; Sánchez-Gracia et al., 2009). Using this framework, we tested several models in which heterogeneous turnover rates were allowed for specific clades according to the respective host ranges. The *Mspe* and *Mall* models, which admit different turnover rates for columnar cacti specialists, prickly pear specialists, and fruit-dwellers, were the best fit to the turnover dynamics of this family. As proposed by Oliveira et al. (2012), we assumed that the ancestral host of the *repleta* group are prickly pear cacti (*Opuntia*-like). Our results support the hypothesis that ecological specialization, particularly the acquisition of cactophily, can account, at least in part, for gene turnover dynamics of the OBP family in *Drosophila*. Adaptation to new hosts has been key in the radiation of several groups of species, such as the well-documented cases of Hawaiian *Drosophila* (O'Grady and DeSalle, 2018b), and cactus-dwelling flies of the *repleta* group (Etges, 2019; Hasson et al., 2019; Markow, 2019). In the latter, the case of *D. mojavensis* is worth mentioning since at least three host shifts occurred in the last My (Allan and Matzkin, 2019). Indeed, transcriptomic studies

revealed changes in expression profiles of genes encoding olfactory and GR induced by the use of alternative hosts between subspecies (Matzkin, 2012; Matzkin and Markow, 2013). Likewise, differences in expression profiles of various chemosensory gene families (including OBPs) in the weevils *Eucryptorrhynchus scrobiculatus* and *E. brandti*, that feed on *Ailanthus altissima* (tree of heaven), have been shown to be involved in the recognition of specific sections of the plant, reducing competition on the same host (Wen et al., 2018). These pieces of evidence suggest that expression of chemosensory genes play an important role in the recognition of specific substrates.

Although the functional role and ecological significance of OBPs in insects have been well studied, the evolution of this family remains unclear (Vogt and Riddiford, 1981; Fan et al., 2011; Leal, 2013). Some studies provide examples of expansions or contractions throughout evolution that do not correlate with particular lifestyles or host shifts, impairing the identification of the main forces driving the evolution of this family (He et al., 2016). From a theoretical point of view, expansions may promote functional diversification, allowing OBPs to specialize in the recognition of diverse target odorants (Auer et al., 2021). However, functional overlap among OBPs has been reported in *D. melanogaster* (Larter et al., 2016), suggesting that an increase in the OBP repertoire does not necessarily imply an increase in the spectrum of recognized targets. In fact, several OBPs can distinguish the same or chemically similar targets (Zhang et al., 2011).

Interestingly, OBPs can be involved in other physiological processes such as reproduction and pheromone detection during mating behavior. Moreover, OBPs expression has been detected in organs not related to chemoperception in insects, such as accessory glands and ovaries (Forêt and Maleszka, 2006; Zhou et al., 2009; Shorter et al., 2016). Therefore, changes in the OBP repertoire are expected to pleiotropically affect behavior, host use, and reproduction.

According to our estimates in internal nodes, the ancestral OBP family in the genus *Drosophila* had 44 genes, a size similar to that also found in the *virilis-repleta* radiation. However, to elucidate more accurately the ancestral size of the family, it is necessary to compare the repertoire of *Drosophila* OBPs with outgroup taxa. Estimates of OBP family size in other Diptera showed a wide range of variation both within and between families. For example, 20 OBPs have been described in the hematophagous tsetse-fly *Glossina morsitans* (Glossinidae) (Liu et al., 2010) and 87, in the house fly *Musca domestica* (Muscidae) (Scott et al., 2014), and 69–111, in hematophagous mosquitoes of the family Culicidae (Manoharan et al., 2013; He et al., 2016). Divergence times between these taxa and *Drosophila* go back to more than 250 My ago for mosquitoes and to ~75 My ago for Muscidae and Glossinidae (Wiegmann et al., 2011). All in all, these results point out the heterogeneous evolutionary dynamics of the OBP repertoire within and between Dipteran families.

Even though the range of variation of the OBP repertoire in Drosophilidae is not as wide as in other Dipteran families, we detected intra-specific variation within *D. koepferae*. Recent phylogenomic studies revealed that Bolivian and Argentinian *D. koepferae* diverged 0.3–0.7 My ago according to genomic data (Moreyra et al., 2019, 2022). These currently allopatric populations are distributed in a geographic area that has been exposed to paleoclimatic changes that deeply affected the distribution of cacti. Moreover, *D. koepferae* has been recovered from rotting columnar cacti such as *Trichocereus terscheckii*, *T. candicans*, and *Cereus validus* in Northwestern Argentina, and from *Neoraimondia herzogiana* in southern Bolivia (Fontdevila et al., 1988; Hasson et al., 1992). Lastly, demography can also influence changes in the repertoire of chemosensory families. As Gardiner et al. (2008) reported, endemic *Drosophila* species with small population sizes lose more OR and GR genes. However, to investigate the roles of ecological or demographic factors in the evolution of the OBP repertoire between *D. koepferae* populations must await for population genomic and transcriptomic studies.

Most species of the *repleta* group are cactus dwellers, of which prickly pears (genus *Opuntia*) are the preferred hosts of many species, probably due to the low toxicity as compared to columnar cacti of the subfamily Cactoideae. The latter includes species capable of producing compounds with toxic effects as defenses against herbivores (War et al., 2018), e.g., alkaloids whose concentration increases in decomposing cactus tissues utilized by *Drosophila* (Ganter et al., 2017). Within the *buzzatii* and *mojavensis* clusters, *D. buzzatii* and *D. navojoa* are mainly associated with prickly pears, whereas more derived species are mainly associated with columnar cacti of different genera (Manfrin and Sene, 2006; Matzkin, 2014; Hasson et al., 2019). Transcriptomic studies have shown that genes associated with detoxification are part of complex responses of flies when facing chemically hostile hosts like columnar cacti (Matzkin, 2014; De Panis et al., 2016, 2022; Diaz et al., 2018). Other studies revealed signals of positive selection in detoxifying genes of the Cytochrome P450 family, likely as an evolutionary response to the interaction with toxicologically aggressive hosts (e.g., Matsuo et al., 2007; Hungate et al., 2013). Moreover, Hoang et al. (2015) found that co-expression of genes involved in detoxification and chemoperception, located in contiguous genomic regions, are under strong selective pressure in the cactophilic *D. mettleri*. In this vein, chemosensory and detoxifying proteins can act in concert to protect olfactory receptor neurons against volatile natural toxins or insecticides (Ding and Kaminsky, 2003). For instance, the Cytochrome P450 family is key in the degradation of alkaloids (Wojtasek and Leal, 1999), one of the main chemical weapons produced by columnar cacti; and certain OBPs are capable of kidnapping toxic ligands in the lymph (Steinbrecht, 1998). These pieces of evidence give support to the hypothesis that the evolution of the chemosensory and detoxifying systems evolved in concert during ecological specialization in *Drosophila* (Anholt, 2020).

In summary, our molecular evolutionary analyses show that the main force driving the evolution of the OBP family is negative selection as it has been reported for other chemosensory families (Vieira et al., 2007; Diaz et al., 2018). However, signals of positive selection were found in a set of OBPs. These facts point out the adaptive relevance of OBPs, which presumably have essential roles in the detection of food, hosts, the exploration of new niches, and, likely, reproduction.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: http://ftp.flybase.net/releases/FB2021_03/, NCBI (GCA_001654015.2, GCA_001654025.1, GCA_003285905.2, and GCA_013340165.1).

Author contributions

JJR, JR, JH, and EH contributed to the conception and design of the study. NNM performed the annotation and assembly of *buzzatii* cluster genomes. JJR and VP performed the cluster analysis and wrote scripts for data analysis. JJR wrote the first draft of the manuscript and participated in all stages of the study. NNM, JH, and EH wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

JJR was in part supported by the Carolina Foundation (Spain), through a full grant for a research visit to Universitat de Barcelona and JR's laboratory. This work was supported by grants awarded by the University of Buenos Aires, CONICET, and ANPCyT (Argentina) to EH and the Ministerio de Ciencia, Innovación y Universidades de España, código oficial grant no. PID2019-103947GB-C21.

Acknowledgments

We wish to thank two reviewers for insightful comments and constructive criticisms that helped to improve previous versions of this work. We thank Alejandro Sánchez-Gracia and the Evolutionary Genomics and Bioinformatics Lab at the University of Barcelona Spain, for advice in phylogenetic and molecular evolutionary analyses. JJR and NNM are recipients of doctoral and postdoctoral scholarships awarded by CONICET, respectively. JH and EH are members of Carrera del Investigador of CONICET. Grant of Ministerio de Ciencia, Innovación y Universidades de España, código oficial PID2019-103947GB-C21.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fevo.2022.957247/full#supplementary-material>

References

- Abascal, F., Zardoya, R., and Telford, M. J. (2010). TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res.* 38, W7–W13. doi: 10.1093/nar/gkq291
- Allan, C. W., and Matzkin, L. M. (2019). Genomic analysis of the four ecologically distinct cactus host populations of *Drosophila mojavensis*. *BMC Genomics* 20:732. doi: 10.1186/s12864-019-6097-z
- Almeida, F. C., Sanchez-Gracia, A., Campos, J. L., and Rozas, J. (2014). Family size evolution in *Drosophila* chemosensory gene families: A comparative analysis with a critical appraisal of methods. *Genome Biol. Evol.* 6, 1669–1682. doi: 10.1093/gbe/evu130
- Anholt, R. R. (2020). Chemosensation and evolution of *Drosophila* host plant selection. *iScience* 23:100799. doi: 10.1016/j.isci.2019.100799
- Arguello, J. R., and Benton, R. (2017). Open questions: Tackling Darwin's "instincts": The genetic basis of behavioral evolution. *BMC Biol.* 15:26. doi: 10.1186/s12915-017-0369-3
- Auer, T. O., Shahandeh, M. P., and Benton, R. (2021). *Drosophila sechellia*: A Genetic Model for Behavioral Evolution and Neuroecology. *Annu. Rev. Genet.* 55, 527–554. doi: 10.1146/annurev-genet-071719-020719
- Bhutkar, A., Schaeffer, S. W., Russo, S. M., Xu, M., Smith, T. F., and Gelbart, W. M. (2008). Chromosomal rearrangement inferred from comparisons of 12 *Drosophila* genomes. *Genetics* 179, 1657–1680. doi: 10.1534/genetics.107.086108
- Brand, P., Ramírez, S. R., Leese, F., Quezada-Euan, J. J. G., Tollrian, R., and Eltz, T. (2015). Rapid evolution of chemosensory receptor genes in a pair of sibling species of orchid bees (*Apidae: Euglossini*). *BMC Evol. Biol.* 15:1–16. doi: 10.1186/s12862-015-0451-9
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: Architecture and applications. *BMC Bioinform.* 10:421. doi: 10.1186/1471-2105-10-421
- Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348
- Clark, A. G., Eisen, M. B., Smith, D. R., Bergman, C. M., Oliver, B., Markow, T. A., et al. (2007). Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450, 203–218. doi: 10.1038/nature06341
- Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2011). ProtTest 3: Fast selection of best-fit models of protein evolution. *Bioinformatics* 27, 1164–1165. doi: 10.1093/bioinformatics/btr088
- De Panis, D. N., Padró, J., Furió Tarrí, P., Tarazona, S., Milla Carmona, P. S., Soto, I. M., et al. (2016). Transcriptome modulation during host shift is driven by secondary metabolites in desert *Drosophila*. *Mol. Ecol.* 25, 4534–4550. doi: 10.1111/mec.13785
- De Panis, D., Dopazo, H., Bongcam-Rudloff, E., Conesa, A., and Hasson, E. (2022). Transcriptional responses are oriented towards different components of the rearing environment in two *Drosophila* sibling species. *BMC Genomics* 23:515. doi: 10.1186/s12864-022-08745-9
- Díaz, F., Allan, C. W., and Matzkin, L. M. (2018). Positive selection at sites of chemosensory genes is associated with the recent divergence and local ecological adaptation in cactophilic *Drosophila*. *BMC Evol. Biol.* 18:144. doi: 10.1186/s12862-018-1250-x
- Ding, X., and Kaminsky, L. S. (2003). Human extrahepatic cytochromes P450: Function in xenobiotic metabolism and tissue-selective chemical toxicity in the respiratory and gastrointestinal tracts. *Annu. Rev. Pharmacol. Toxicol.* 43:149. doi: 10.1146/annurev.pharmtox.43.100901.140251
- Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* 7:e1002195. doi: 10.1371/journal.pcbi.1002195
- Emms, D. M., and Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20:238. doi: 10.1186/s13059-019-1832-y
- Etges, W. J. (2019). Evolutionary genomics of host plant adaptation: Insights from *Drosophila*. *Curr. Opin. Insect Sci.* 36, 96–102. doi: 10.1016/j.cois.2019.08.011
- Eyun, S. I., Soh, H. Y., Posavi, M., Munro, J. B., Hughes, D. S., Murali, S. C., et al. (2017). Evolutionary history of chemosensory-related gene families across the Arthropoda. *Mol. Biol. Evol.* 34, 1838–1862. doi: 10.1093/molbev/msx147
- Fan, J., Francis, F., Liu, Y., Chen, J., and Cheng, D. (2011). An overview of odorant-binding protein functions in insect peripheral olfactory reception. *Genet. Mol. Res.* 10, 3056–3069. doi: 10.4238/2011.December.8.2
- Fontdevila, A., Pla, C., Hasson, E., Naveira, H., and Ruiz, A. (1988). *Drosophila* koepferae: A New Member of the *Drosophila* serido (*Diptera: Drosophilidae*) Superspecies Taxon. *Ann. Entomol. Soc. Am.* 81, 380–385. doi: 10.1093/aesa/81.3.380
- Forêt, S., and Maleszka, R. (2006). Function and evolution of a gene family encoding odorant-binding-like proteins in a social insect, the honey bee (*Apis mellifera*). *Genome Res.* 16, 1404–1413. doi: 10.1101/gr.5075706
- Forister, M. L., Dyer, L. A., Singer, M. S., Stireman, J. O. III, and Lill, J. T. (2012). Revisiting the evolution of ecological specialization, with emphasis on insect–plant interactions. *Ecology* 93, 981–991. doi: 10.1890/11-0650.1
- Ganter, P. F., Morais, P. B., and Rosa, C. A. (2017). “Yeasts in cacti and tropical fruit,” in *Yeasts in Natural Ecosystems: Diversity*, (eds) P. Buzzini, M.-A. Lachance, A. Yurkov (Cham: Springer), 225–264. doi: 10.1007/978-3-319-62683-3_8
- Gardiner, A., Barker, D., Butlin, R. K., Jordan, W. C., and Ritchie, M. G. (2008). *Drosophila* chemoreceptor gene evolution: Selection, specialization and genome size. *Mol. Ecol.* 17, 1648–1657. doi: 10.1111/j.1365-294X.2008.03713.x
- Goldman-Huertas, B., Mitchell, R. F., Lapoint, R. T., Faucher, C. P., Hildebrand, J. G., and Whiteman, N. K. (2015). Evolution of herbivory in *Drosophilidae* linked to loss of behaviors, antennal responses, odorant receptors, and ancestral diet. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3026–3031. doi: 10.1073/pnas.1424656112
- Grimaldi, D., and Agosti, D. (2000). A formicine in New Jersey Cretaceous amber (*Hymenoptera: Formicidae*) and early evolution of the ants. *Proc. Natl. Acad. Sci. U.S.A.* 97, 13678–13683. doi: 10.1073/pnas.240452097
- Grimaldi, D., and Engel, M. S. (2005). *Evolution of the Insects*. New York, NY: Cambridge University Press.
- Guillén, Y., Rius, N., Delprat, A., Williford, A., Muiyas, F., Puig, M., et al. (2014). Genomics of ecological adaptation in cactophilic *Drosophila*. *Genome Biol. Evol.* 7, 349–366. doi: 10.1093/gbe/evu291
- Hahn, M. W., De Bie, T., Stajich, J. E., Nguyen, C., and Cristianini, N. (2005). Estimating the tempo and mode of gene family evolution from comparative genomic data. *Genome Res.* 15, 1153–1160. doi: 10.1101/gr.3567505

- Hasson, E., De Panis, D., Hurtado, J., and Mensch, J. (2019). Host plant adaptation in cactophilic species of the *Drosophila buzzatii* cluster: Fitness and transcriptomics. *J. Hered.* 110, 46–57. doi: 10.1093/jhered/esy043
- Hasson, E., Naveira, H., and Fontdevila, A. (1992). The breeding sites of Argentinian cactophilic species of the *Drosophila mulleri* complex (subgenus *Drosophila-repleta* group). *Rev. Chil. Hist. Nat.* 65, 319–326.
- He, X., He, Z. B., Zhang, Y. J., Zhou, Y., Xian, P. J., Qiao, L., et al. (2016). Genome-wide identification and characterization of odorant-binding protein (OBP) genes in the malaria vector *Anopheles sinensis* (Diptera: Culicidae). *Insect Sci.* 23, 366–376. doi: 10.1111/1744-7917.12333
- Hildebrand, J. G., and Shepherd, G. M. (1997). Mechanisms of olfactory discrimination: Converging evidence for common principles across phyla. *Annu. Rev. Neurosci.* 20, 595–631. doi: 10.1146/annurev.neuro.20.1.595
- Hoang, K., Matzkin, L. M., and Bono, J. M. (2015). Transcriptional variation associated with cactus host plant adaptation in *Drosophila mettleri* populations. *Mol. Ecol.* 24, 5186–5199. doi: 10.1111/mec.13388
- Hungate, E. A., Earley, E. J., Boussy, I. A., Turissini, D. A., Ting, C. T., Moran, J. R., et al. (2013). A locus in *Drosophila sechellia* affecting tolerance of a host plant toxin. *Genetics* 195, 1063–1075. doi: 10.1534/genetics.113.154773
- Hurtado, J. P., Almeida, F., Revale, S., and Hasson, E. (2019). Revised phylogenetic relationships within the *Drosophila buzzatii* species cluster (Diptera: Drosophilidae: *Drosophila repleta* group) using genomic data. *Arthropod. Syst. Phylogeny.* 77, 239–250. doi: 10.26049/ASP77-2-2019-03
- Jezovit, J. A., Levine, J. D., and Schneider, J. (2017). Phylogeny, environment and sexual communication across the *Drosophila* genus. *J. Exp. Biol.* 220, 42–52. doi: 10.1242/jeb.143008
- Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031
- Joseph, R. M., and Carlson, J. R. (2015). *Drosophila* chemoreceptors: A molecular interface between the chemical world and the brain. *Trends Genet.* 31, 683–695. doi: 10.1016/j.tig.2015.09.005
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kaupp, U. (2010). Olfactory signalling in vertebrates and insects: Differences and commonalities. *Nat. Rev. Neurosci.* 11, 188–200. doi: 10.1038/nrn2789
- Kim, B. Y., Wang, J. R., Miller, D. E., Barmina, O., Delaney, E., Thompson, A., et al. (2021). Highly contiguous assemblies of 101 drosophilid genomes. *Elife* 10:e66405. doi: 10.7554/eLife.66405
- Kosakovsky Pond, S. L., Poon, A. F., Velazquez, R., Weaver, S., Hepler, N. L., Murrell, B., et al. (2020). HyPhy 2.5—a customizable platform for evolutionary hypothesis testing using phylogenies. *Mol. Biol. Evol.* 37, 295–299. doi: 10.1093/molbev/msz197
- Krieger, M. J., and Ross, K. G. (2002). Identification of a major gene regulating complex social behavior. *Science* 295, 328–332. doi: 10.1126/science.1065247
- Kück, P., and Meusemann, K. (2010). FASconCAT: Convenient handling of data matrices. *Mol. Phylog. Evol.* 56, 1115–1118. doi: 10.1016/j.ympev.2010.04.024
- Kulmuni, J., Wurm, Y., and Pamilo, P. (2013). Comparative genomics of chemosensory protein genes reveals rapid evolution and positive selection in ant-specific duplicates. *Heredity* 110, 538–547. doi: 10.1038/hdy.2012.122
- Kumar, S., Stecher, G., Suleski, M., and Hedges, S. B. (2017). TimeTree: A resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* 34, 1812–1819. doi: 10.1093/molbev/msx116
- Larkin, A., Marygold, S. J., Antonazzo, G., Attrill, H., dos Santos, G., Garapati, P. V., et al. (2021). FlyBase: Updates to the *Drosophila melanogaster* knowledge base. *Nucleic Acids Res.* 49, D899–D907. doi: 10.1093/nar/gkaa1026
- Larter, N. K., Sun, J. S., and Carlson, J. R. (2016). Organization and function of *Drosophila* odorant binding proteins. *Elife* 5:e20242. doi: 10.7554/eLife.20242.001
- Leal, W. S. (2013). Odorant reception in insects: Roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* 58, 373–391. doi: 10.1146/annurev-ento-120811-153635
- Lee, E., Harris, N., Gibson, M., Chetty, R., and Lewis, S. (2009). Apollo: A community resource for genome annotation editing. *Bioinformatics* 25, 1836–1837. doi: 10.1093/bioinformatics/btp314
- Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi: 10.1093/nar/gkab301
- Librado, P., and Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187
- Librado, P., and Rozas, J. (2013). Uncovering the functional constraints underlying the genomic organization of the odorant-binding protein genes. *Genome Biol. Evol.* 5, 2096–2108. doi: 10.1093/gbe/evt158
- Librado, P., Vieira, F. G., and Rozas, J. (2012). BadiRate: Estimating family turnover rates by likelihood-based methods. *Bioinformatics* 28, 279–281. doi: 10.1093/bioinformatics/btr623
- Liu, R., Lehane, S., He, X., Lehane, M., Hertz-Fowler, C., Berriman, M., et al. (2010). Characterisations of odorant-binding proteins in the tsetse fly *Glossina morsitans morsitans*. *Cell Mol. Life Sci.* 67, 919–929. doi: 10.1007/s00018-009-0221-1
- Manfrin, M. H., and Sene, F. M. (2006). Cactophilic *Drosophila* in South America: A model for evolutionary studies. *Genetica* 126, 57–75. doi: 10.1007/s10709-005-1432-5
- Manoharan, M., Ng Fuk, Chong, M., Vaïtinadapoulé, A., Frumence, E., Sowdhamini, R., et al. (2013). Comparative genomics of odorant binding proteins in *Anopheles gambiae*, *Aedes aegypti*, and *Culex quinquefasciatus*. *Genome Biol. Evol.* 5, 163–180. doi: 10.1093/gbe/evs131
- Marchler-Bauer, A., Anderson, J. B., Cherukuri, P. F., DeWeese-Scott, C., Geer, L. Y., Gwadz, M., et al. (2005). CDD: A Conserved Domain Database for protein classification. *Nucleic Acids Res.* 33, D192–D196. doi: 10.1093/nar/gki069
- Markow, T. A. (2019). Host use and host shifts in *Drosophila*. *Curr. Opin. Insect Sci.* 31, 139–145. doi: 10.1016/j.cois.2019.01.006
- Matsuo, T., Sugaya, S., Yasukawa, J., Aigaki, T., and Fuyama, Y. (2007). Odorant binding proteins OBP57d and OBP57e affect taste perception and host-plant preference in *Drosophila sechellia*. *PLoS Biol.* 5:e118. doi: 10.1371/journal.pbio.0050118
- Matzkin, L. M. (2012). Population transcriptomics of cactus host shifts in *Drosophila mojavensis*. *Mol. Ecol.* 21, 2428–2439. doi: 10.1111/j.1365-294X.2012.05549.x
- Matzkin, L. M. (2014). Ecological genomics of host shifts in *Drosophila mojavensis*. *Ecol. Genomics* 781, 233–247. doi: 10.1007/978-94-007-7347-9_12
- Matzkin, L. M., and Markow, T. A. (2013). “Transcriptional differentiation across the four cactus host races of *Drosophila mojavensis*,” in *Speciation: Natural Processes, Genetics and Biodiversity*, ed. P. Michalak (New York, NY: Nova Scientific Publishers), 119–135. doi: 10.1007/978-94-007-7347-9_12
- McBride, C. S. (2007). Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*. *Proc. Nat. Acad. Sci. U.S.A.* 104, 4996–5001. doi: 10.1073/pnas.0608424104
- McBride, C. S., and Arguello, J. R. (2007). Five *Drosophila* genomes reveal nonneutral evolution and the signature of host specialization in the chemoreceptor superfamily. *Genetics* 177, 1395–1416. doi: 10.1534/genetics.107.078683
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L., et al. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Res.* 49, D412–D419. doi: 10.1093/nar/gkaa913
- Moreyra, N. N., Cunha Almeida, F., Allan, C., Frankel, N., Matzkin, L., and Hasson, E. (2022). Phylogenomics provides insights into the evolution of cactophily and host plant shifts adaptation in *Drosophila*. *eLife* [Preprint] doi: 10.1101/2022.04.29.490106
- Moreyra, N. N., Mensch, J., Hurtado, J., Almeida, F., Laprida, C., and Hasson, E. (2019). What does mitogenomics tell us about the evolutionary history of the *Drosophila buzzatii* cluster (repleta group)? *PLoS One* 14:e0220676. doi: 10.1371/journal.pone.0220676
- Murrell, B., Weaver, S., Smith, M. D., Wertheim, J. O., Murrell, S., and Aylward, A. (2015). Gene-wide identification of episodic selection. *Mol. Biol. Evol.* 32, 1365–1371. doi: 10.1093/molbev/msv035
- Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., and Kosakovsky Pond, S. L. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Gene.* 8:e1002764. doi: 10.1371/journal.pgen.1002764
- Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- O’Grady, P. M., and DeSalle, R. (2018a). Phylogeny of the genus *Drosophila*. *Genetics* 209, 1–25. doi: 10.1534/genetics.117.300583
- O’Grady, P., and DeSalle, R. (2018b). Hawaiian *Drosophila* as an evolutionary model clade: Days of future past. *Bioessays* 40:1700246. doi: 10.1002/bies.201700246
- Oliveira, D. C., Almeida, F. C., O’Grady, P. M., Armella, M. A., DeSalle, R., and Etges, W. J. (2012). Monophyly, divergence times, and evolution of host plant use inferred from a revised phylogeny of the *Drosophila repleta* species group. *Mol. Phylogenet. Evol.* 64, 533–544. doi: 10.1016/j.ympev.2012.05.012

- Pelosi, P. (1994). Odorant-binding proteins. *Crit. Rev. Biochem. Mol. Biol.* 29, 199–228. doi: 10.3109/10409239409086801
- Pelosi, P. (1996). Perireceptor events in olfaction. *J. Neurobiol.* 30, 3–19. doi: 10.1002/(SICI)1097-4695
- Rane, R. V., Pearce, S. L., Li, F., Coppin, C., Schiffer, M., Shirriffs, J., et al. (2019). Genomic changes associated with adaptation to arid environments in cactophilic *Drosophila* species. *BMC Genomics* 20:52. doi: 10.1186/s12864-018-5413-3
- Ranz, J. M., Casals, F., and Ruiz, A. (2001). How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. *Genome Res.* 11, 230–239. doi: 10.1101/gr.162901
- Ranz, J. M., Segarra, C., and Ruiz, A. (1997). Chromosomal homology and molecular organization of Muller's elements D and E in the *Drosophila repleta* species group. *Genetics* 145, 281–295. doi: 10.1093/genetics/145.2.281
- Richards, S., Liu, Y., Bettencourt, B. R., Hradscky, P., Letovsky, S., Nielsen, R., et al. (2005). Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and cis-element evolution. *Genome Res.* 15, 1–18. doi: 10.1101/gr.3059305
- Robertson, H. M. (2019). Molecular evolution of the major arthropod chemoreceptor gene families. *Annu. Rev. Entomol.* 64, 227–242. doi: 10.1146/annurev-ento-020117-043322
- Sanchez-Flores, A., Peñaloza, F., Carpinteyro-Ponce, J., Nazario-Yepiz, N., Abreu-Goodger, C., Machado, C. A., Markow, T. A. (2016). Genome evolution in three species of cactophilic *Drosophila*. *G3, Genes, Genomes. Genetics* 6, 3097–3105 doi: 10.1534/g3.116.033779
- Sánchez-Gracia, A., Vieira, F. G., and Rozas, J. (2009). Molecular evolution of the major chemosensory gene families in insects. *Heredity* 103, 208–216. doi: 10.1038/hdy.2009.55
- Sanderson, M. J. (2003). r8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19, 301–302. doi: 10.1093/bioinformatics/19.2.301
- Sayers, E. W., Cavanaugh, M., Clark, K., Pruitt, K. D., Schoch, C. L., Sherry, S. T., et al. (2021). GenBank. *Nucleic Acids Res.* 49, D92–D96. doi: 10.1093/nar/gkv1276
- Schaeffer, S. W. (2018). Muller “elements” in *Drosophila*: How the search for the genetic basis for speciation led to the birth of comparative genomics. *Genetics* 210, 3–13. doi: 10.1534/genetics.118.301084
- Scott, J. G., Warren, W. C., Beukeboom, L. W., Bopp, D., Clark, A. G., Giers, S. D., et al. (2014). Genome of the house fly, *Musca domestica* L., a global vector of diseases with adaptations to a septic environment. *Genome Biol.* 15, 1–17. doi: 10.1186/s13059-014-0466-3
- Shorter, J. R., Dembeck, L. M., Everett, L. J., Morozova, T. V., Arya, G. H., Turlapati, L., et al. (2016). Obp56h modulates mating behavior in *Drosophila melanogaster*. *G3* 6, 3335–3342. doi: 10.1534/g3.116.034595
- Silvestro, D., Bacon, C. D., Ding, W., Zhang, Q., Donoghue, P. C., Antonelli, A., et al. (2021). Fossil data support a pre-Cretaceous origin of flowering plants. *Nat. Ecol. Evol.* 5, 449–457. doi: 10.1038/s41559-020-01387-8
- Smith, M. D., Wertheim, J. O., Weaver, S., Murrell, B., Scheffler, K., and Kosakovsky Pond, S. L. (2015). Less is more: An adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* 32, 1342–1353. doi: 10.1093/molbev/msv022
- Stamatakis, A. (2014). RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Steinbrecht, R. A. (1998). Odorant-binding proteins: Expression and function. *Ann. N.Y. Acad. Sci.* 855, 323–332. doi: 10.1111/j.1749-6632.1998.tb10591.x
- van der Niet, T., and Johnson, S. D. (2012). Phylogenetic evidence for pollinator-driven diversification of angiosperms. *Trends Ecol. Evol.* 27, 353–361. doi: 10.1016/j.tree.2012.02.002
- Vieira, F. G., and 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 Biol. Evol.* 3, 476–490. doi: 10.1093/gbe/evr033
- Vieira, F. G., Sánchez-Gracia, A., and Rozas, J. (2007). Comparative genomic analysis of the odorant-binding protein family in 12 *Drosophila* genomes: Purifying selection and birth-and-death evolution. *Genome Biol.* 8:R235. doi: 10.1186/gb-2007-8-11-r235
- Vizueta, J., Escuer, P., Frías-López, C., Guirao-Rico, S., Hering, L., Mayer, G., et al. (2020b). Evolutionary history of major chemosensory gene families across Panarthropoda. *Mol. Biol. Evol.* 37, 3601–3615. doi: 10.1093/molbev/msaa197
- Vizueta, J., Escuer, P., Sánchez-Gracia, A., and Rozas, J. (2020a). Genome mining and sequence analysis of chemosensory soluble proteins in arthropods. *Methods Enzymol.* 642, 1–20. doi: 10.1016/bs.mie.2020.05.015
- Vizueta, J., Frías-López, C., Macías-Hernández, N., Arnedo, M. A., Sánchez-Gracia, A., and 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. doi: 10.1093/gbe/evw296
- Vizueta, J., Sánchez-Gracia, A., and Rozas, J. (2020c). BITACORA: A comprehensive tool for the identification and annotation of gene families in genome assemblies. *Mol. Ecol. Resour.* 20, 1445–1452. doi: 10.1111/1755-0998.13202
- Vogt, R. G., and Riddiford, L. M. (1981). Pheromone binding and inactivation by moth antennae. *Nature* 293, 161–163. doi: 10.1038/293161a0
- Wang, B., Xu, C., and Jarzembowski, E. A. (2022). Ecological radiations of insects in the Mesozoic. *Trends Ecol. Evol.* 37, 529–540. doi: 10.1016/j.tree.2022.02.007
- War, A. R., Taggar, G. K., Hussain, B., Taggar, M. S., Nair, R. M., and Sharma, H. C. (2018). Plant defence against herbivory and insect adaptations. *AoB Plants* 10:ly037. doi: 10.1093/aobpla/ply037
- Weisman, C., Murray, A. W., and Eddy, S. R. (2022). Mixing genome annotation methods in a comparative analysis inflates the apparent number of lineage-specific genes. *Curr Biol.* 32, 2632–2639.e2. doi: 10.1016/j.cub.2022.04.085
- Wen, X., Wang, Q., Gao, P., and Wen, J. (2018). Identification and comparison of chemosensory genes in the antennal transcriptomes of *Eucryptorrhynchus scrobiculatus* and *E. brandti* fed on *Ailanthus altissima*. *Front. Physiol.* 9:1652. doi: 10.3389/fphys.2018.01652
- Wiegmann, B. M., Trautweina, M. D., Winklera, I. S., Barr, N. B., et al. (2011). Episodic radiations in the fly tree of life. *Proc. Natl. Acad. Sci. U.S.A.* 108, 5690–5695. doi: 10.1073/pnas.1012675108
- Wojtasek, H., and Leal, W. S. (1999). Degradation of an alkaloid pheromone from the pale-brown chafer, *Phyllopertha diversa* (Coleoptera: Scarabaeidae), by an insect olfactory cytochrome P450. *FEBS Lett.* 458, 333–336. doi: 10.1016/S0014-5793(99)01178-3
- Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088
- Yi, G., Sze, S. H., and Thon, M. R. (2007). Identifying clusters of functionally related genes in genomes. *Bioinformatics* 23, 1053–1060. doi: 10.1093/bioinformatics/btl673
- Zhang, S., Chen, L. Z., Gu, S. H., Cui, J. J., Gao, X. W., Zhang, Y. J., et al. (2011). Binding characterization of recombinant odorant-binding proteins from the parasitic wasp, *Microplitis mediator* (Hymenoptera: Braconidae). *J. Chem. Ecol.* 37, 189–194. doi: 10.1007/s10886-010-9902-3
- Zheng, D., Chang, S. C., Perrichot, V., Dutta, S., Rudra, A., Mu, L., et al. (2018). A Late Cretaceous amber biota from central Myanmar. *Nat. Commun.* 9, 1–6. doi: 10.1038/s41467-018-05650-2
- Zhou, S., Stone, E. A., Mackay, T. F., and Anholt, R. R. (2009). Plasticity of the chemoreceptor repertoire in *Drosophila melanogaster*. *PLoS Genet.* 5:e1000681. doi: 10.1371/journal.pgen.1000681