Expansion and Accelerated Evolution of 9-Exon Odorant Receptors in *Polistes* Paper Wasps

Andrew W. Legan,* Christopher M. Jernigan, Sara E. Miller, Matthieu F. Fuchs, and Michael J. Sheehan*

Laboratory for Animal Social Evolution and Recognition, Department of Neurobiology and Behavior, Cornell University, Ithaca, NY 14853, USA

***Corresponding authors:** E-mails: awl75@cornell.edu; msheehan@cornell.edu

Associate editor: Deepa Agashe

Abstract

Independent origins of sociality in bees and ants are associated with independent expansions of particular odorant receptor (OR) gene subfamilies. In ants, one clade within the OR gene family, the 9-exon subfamily, has dramatically expanded. These receptors detect cuticular hydrocarbons (CHCs), key social signaling molecules in insects. It is unclear to what extent 9-exon OR subfamily expansion is associated with the independent evolution of sociality across Hymenoptera, warranting studies of taxa with independently derived social behavior. Here, we describe OR gene family evolution in the northern paper wasp, *Polistes fuscatus*, and compare it to four additional paper wasp species spanning ~40 million years of evolutionary divergence. We find 200 putatively functional OR genes in *P. fuscatus*, matching predictions from neuroanatomy, and more than half of these are in the 9-exon subfamily. Most OR gene expansions are tandemly arrayed at orthologous loci in *Polistes* genomes, and microsynteny analysis shows species-specific gain and loss of 9-exon ORs within tandem arrays. There is evidence of episodic positive diversifying selection shaping ORs in expanded subfamilies. Values of omega (d_N/d_S) are higher among 9-exon ORs compared to other OR subfamilies. Within the *Polistes* OR gene tree, branches in the 9-exon OR clade experience relaxed negative (relaxed purifying) selection relative to other branches in the tree. Patterns of OR evolution within *Polistes* are consistent with 9-exon OR function in CHC perception by combinatorial coding, with both natural selection and neutral drift contributing to interspecies differences in gene copy number and sequence.

Key words: **birth-and-death process, comparative genomics, tandem array, olfaction, social insect, antennal lobe glomeruli.**

Article

Introduction

Odorant/olfactory receptors (ORs) are among the largest gene families in animal genomes, and variation in the OR repertoire is hypothesized to reflect aspects of species chemosensory ecology. From the standpoint of molecular evolution, the ORs of insects and mammals have been widely studied as a model to understand the dynamics of gene family evolution (Young et al. 2002; Robertson et al. 2003; Nozawa and Nei 2007; Eirín-López et al. 2012; Nei 2013; Benton 2015; McKenzie and Kronauer 2018). Yet fundamental features of odorant receptor (OR) evolution remain unclear—why do some groups show predominantly conserved OR repertoires across species while others show rapid turnover in gene content or accelerated rates of evolution? Moreover, the relative importance of social interactions, sexual selection, and ecology in shaping patterns of OR evolution within and between clades is poorly understood. Comparative studies of distantly related species have provided insights into the evolutionary processes shaping the insect and mammal OR gene families at broad phylogenetic scales, where there is relatively little 1:1 orthology of receptors among species (Tsutsui 2013; Roux et al. 2014; Freeman et al. 2020; Yan et al. 2020). At the same time, studies within species and between closely related species can reveal the dynamics of receptor evolution at finer timescales and elucidate the process of gene family turnover, as evidenced by studies of insects and mammals (Guo and Kim 2007; McBride et al. 2014; Brand et al. 2015; Karpe et al. 2016; Brand and Ramírez 2017; Cohanim et al. 2018; Miller CH et al. 2020). Recent efforts to sequence a growing number of social insect genomes have suggested that social evolution is associated with expansions within the OR gene family, and the 9-exon OR subfamily in particular has experienced increased gene turnover and sequence evolution relative to other OR subfamilies (Zhou et al. 2012, 2015; LeBoeuf et al. 2013; Engsontia et al. 2015; Kapheim et al. 2015; Karpe et al. 2016, 2017; McKenzie et al. 2016; Saad et al. 2018). Given the importance of olfaction for social insect behavioral ecology, ORs provide a key route to linking genes to diverse and complex behaviors among ants, bees, and wasps. However, there are two major gaps in our knowledge of OR evolution in social insects. First, the hypothesis that social evolution is associated with OR expansions has yet to be tested in all of the independent origins of sociality among Hymenoptera. The independent origins of sociality in wasps provide an

opportunity to compare patterns of OR gene family evolution to those that have been observed within social bees and ants (Hines et al. 2007). Second, the fine-scale dynamics of OR gene family turnover between social insect species remain understudied, since most studies have focused on comparisons between genera or families. A better understanding of the short-term mechanisms of OR evolution provides additional insights into the molecular evolutionary dynamics shaping receptor diversity across more distantly related taxa. The recent release of five genomes of *Polistes* paper wasps spanning ~1–40 million years of divergence provides an opportunity to fill these gaps in our knowledge.

Organisms use chemoreceptor proteins to detect stimuli and provide input to neural circuits that regulate decision-making (Su et al. 2009; Yapici et al. 2014). The insect OR gene family evolved in the ancestor of all insects and constitutes the largest among the insect chemoreceptor gene families, which also include gustatory receptors and ionotropic receptors (Hansson and Stensmyr 2011; Suh et al. 2014; Brand et al. 2018; Fleischer et al. 2018; Vizueta et al. 2020). OR genes code for subunits of heterotetrameric ligand-gated ion channels embedded in the membranes of olfactory receptor neurons (ORNs) (Sato et al. 2008; Butterwick et al. 2018). The odorant receptor coreceptor (Orco) gene codes for a component of all OR complexes and is ubiquitously expressed in olfactory tissue and highly conserved across insects (Fleischer et al. 2018). In addition to expressing Orco, each ORN generally expresses only one OR gene, and the specificity with which an OR binds chemical compounds or families of compounds (ligands) determines the response spectrum of the ORN in which it is expressed (Hallem et al. 2004). At the molecular level, the ORN responses are dependent upon a complex interaction of OR, ligand concentration, and odorant-binding proteins (Vogt et al. 1991; Hallem et al. 2004; Hallem and Carlson 2006; Stensmyr et al. 2012; Mathew et al. 2013; Dweck et al. 2015; Ebrahim et al. 2015; Münch and Galizia 2016). In flies, ORNs expressing a particular OR project to a common glomerulus in the antennal lobe (AL), which generally appears to be true across insects (Couto et al. 2005; Su et al. 2009; Galizia and Rössler 2010).

The molecular evolution of the OR gene family is best described as a birth-and-death process, in which genes are duplicated and deleted over evolutionary time (Nei 2007; Nozawa and Nei 2007; Eirín-López et al. 2012). Both random drift and natural selection are present during this process, determining the extent of OR gene copy number variation and the rate of gene sequence evolution (Nei 2007; Nozawa and Nei 2007). First identified in the common fruit fly *Drosophila melanogaster* and the African malaria mosquito *Anopheles gambiae*, much of our understanding of the relationship between insect OR function and evolution comes from studies of Diptera, which show that the OR gene family can be conserved among species in a genus, with exceptions arising when chemosensory landscapes differ between species (Vosshall et al. 1999; Fox et al. 2001; Robertson et al. 2003). There is prevalent negative (purifying) selection conserving ORs across *Drosophila* species, and the majority of *D. melanogaster* ORs form simple orthologous relationships across the genus (Clark et al. 2007; Guo and Kim 2007; McBride and Arguello 2007; Nozawa and Nei 2007; Sánchez-Gracia et al. 2009; Mansourian and Stensmyr 2015). However, there is evidence of gene loss and accelerated evolution of some ORs during the evolution of host specialization and herbivory in Drosophilids (McBride 2007; McBride and Arguello 2007; Goldman-Huertas et al. 2015). Between genera, the OR repertoire is more variable. The *Aedes aegypti* OR repertoire includes about 131 genes organized in clades that are largely divergent from *Anopheles gambiae*'s 79 OR genes (Fox et al. 2001; Hill et al. 2002; Bohbot et al. 2007). Between Dipteran families, there is considerable variation in OR sequences and copy number, indicating that OR evolution is more dynamic at this phylogenetic scale (Fox et al. 2001; Bohbot et al. 2007; Carey et al. 2010).

The molecular evolution of the OR gene family is dynamic among Hymenoptera genera, with prevalent lineage-specific gene expansions and losses, especially in the 9-exon OR subfamily (Engsontia et al. 2015; Zhou et al. 2015; McKenzie and Kronauer 2018). The 9-exon ORs constitute about one-third of all ant ORs, and have evolved rapidly in ants, leading researchers to propose that 9-exon ORs facilitate recognition of cuticular hydrocarbons (CHCs) (Smith CR et al. 2011; Smith CD et al. 2011; Zhou et al. 2012, 2015; Engsontia et al. 2015; McKenzie et al. 2016). CHCs are used by insects to waterproof the cuticle and to communicate with conspecifics (Blomquist and Bagnères 2010). While less pronounced than in ants, dynamic evolution is also characteristic of 9-exon OR evolution in social bees, which rely on CHCs in communication (Sadd et al. 2015; Karpe et al. 2016, 2017). Functional studies in which ORs were transfected into an empty *D. melanogaster* ORN have verified that at least some 9-exon ORs of the ant *Harpegnathos saltator* overlap in their responses to ligands, with multiple 9-exon ORs responding to the same CHC molecule and unique 9-exon ORs responding to multiple different CHC molecules (Pask et al. 2017; Slone et al. 2017). This is characteristic of combinatorial coding: the process of combining input from multiple ORs that bind overlapping sets of ligands in order to discriminate a larger variety of odors (Malnic et al. 1999; Touhara and Vosshall 2009). Functional ORs are necessary for normal nesting behavior and for nestmate recognition in ants, a process which involves detecting variation in the CHCs on the cuticles of conspecifics (Lavine et al. 1990; van Zweden and d'Ettorre 2010; Sturgis and Gordon 2012; Trible et al. 2017; Yan et al. 2017; Ferguson et al. 2020). Together these studies suggest that 9-exon ORs function in combinatorial coding of CHC perception in ants and potentially in general across Hymenoptera.

Like other Hymenopterans, vespid wasps, including the genus *Polistes*, use CHCs in complex social behaviors (Gamboa et al. 1986, 1996; Dani and Turillazzi 2018). *Polistes* use chemicals as signals and cues during mate attraction, mate compatibility recognition, queen recognition, dominance/fertility signaling, and nestmate recognition (Post and Jeanne 1984; Reed and Landolt 1990; Espelie et al. 1994; Sledge et al. 2001a, 2001b, 2004; Dapporto et al. 2007; Jandt et al. 2014; Oi et al. 2019). The molecular mechanistic

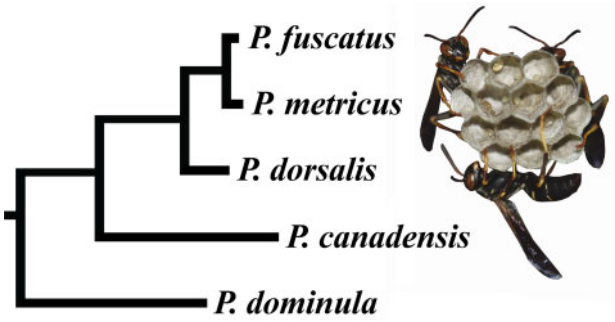

FIG. 1. Phylogeny of five *Polistes* species considered in this study: *P. fuscatus*, *P. metricus*, *P. dorsalis*, *P. canadensis*, and *P. dominula*. The photo to the right of the phylogeny shows *P. fuscatus* foundresses on a nest. Phylogenetic tree based on the 16S ribosomal RNA gene and the cytochrome oxidase subunit I gene.

basis of chemical signal perception in *Polistes* has not been explored, but the importance of olfaction in mediating social behaviors is expected to favor increased copy number of OR genes encoding chemical signal receptors. Neuroanatomical analyses of the antennal lobe of social wasps suggest they possess expanded OR repertoires. In the clonal raider ant, *Ooceraea biroi*, the T6 cluster of the antennal lobe receives inputs from OR neurons expressing 9-exon subfamily ORs in sensilla basiconica (McKenzie et al. 2016). The large T_B cluster in the antennal lobe of Vespid wasps is homologous to the T6 cluster in ants, suggesting the 9-exon subfamily of ORs has also expanded in wasps (Masson and Strambi 1977; Couto et al. 2016, 2017). Recent efforts to sequence *Polistes* genomes provide an opportunity to resolve patterns of OR evolution among closely related species as an independent test of 9-exon OR gene subfamily expansion during social evolution (Patalano et al. 2015; Standage et al. 2016; Miller et al. 2020). We annotated the OR repertoires of five *Polistes* species representing ~40 million years of evolution: *P. fuscatus*, *P. metricus*, *P. dorsalis*, *P. canadensis*, and *P. dominula* (fig. 1). Combining neuroanatomy, manual gene annotation, and molecular evolution analysis, we examined the evolution of *Polistes* ORs with a focus on the 9-exon subfamily. We discover that social wasps, like ants, have an expanded set of 9-exon ORs. Between *Polistes* species, 9-exon ORs exhibit dynamic evolution and relaxed negative selection relative to ORs in other subfamilies, which are highly conserved. Patterns of molecular evolution of the 9-exon OR subfamily in social wasps are consistent with a unique function in combinatorial coding perception of CHCs.

Results

Antennal Lobe Neuroanatomy and Manual Gene Annotation Predict 200 ORs in *P. fuscatus*

In order to predict the OR repertoires of *P. fuscatus* and four other *Polistes* species, we combined fluorescent confocal microscopy of the *P. fuscatus* antennal lobe with manual genome annotation informed by antennal RNAseq. We found 229 glomeruli in the antennal lobe of an adult gyne (female reproductive) (supplementary fig. S1, Supplementary Material online). Across a sample of insects, the number of

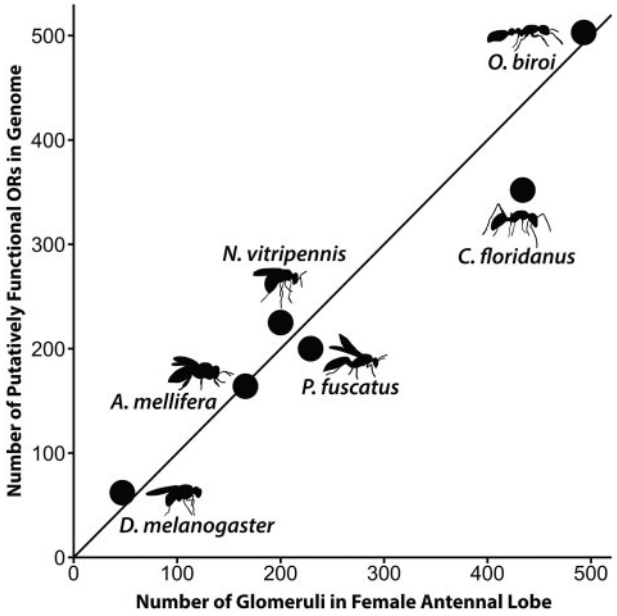

FIG. 2. The number of functional ORs is correlated with the number of antennal lobe glomeruli across insect species (50 glomeruli and 62 ORs in the genome of the common fruit fly *D. melanogaster*, Fishilevich and Vosshall 2005; 166 glomeruli in the worker and 163 intact ORs in the genome of the honey bee *A. mellifera*, Arnold et al. 1985, Robertson and Wanner 2006; ~200 glomeruli in females and 225 intact ORs in the genome of the parasitic wasp *N. vitripennis*, Groothuis et al. 2019, Robertson et al. 2010; ~434 glomeruli in the worker and 352 functional ORs in the genome of the ant *C. floridanus*, Zube and Rössler 2008, Zhou et al. 2012; 493 glomeruli in the worker and 503 intact ORs in the genome of the ant *O. biroi*, McKenzie et al. 2016; McKenzie and Kronauer 2018). The diagonal line represents a line of equality with slope of 1.

intact OR genes in the genome correlates with the number of glomeruli in the antennal lobe, predicting 229 ORs in the *P. fuscatus* genome (fig. 2). Here, we focus on the *P. fuscatus* genome because it has nearly chromosome level scaffolds and is the best assembled *Polistes* genome (Patalano et al. 2015; Standage et al. 2016; Miller et al. 2020) (supplementary table S1, Supplementary Material online). Automated annotation using the MAKER pipeline (Holt and Yandell 2011) without guidance from antennal mRNA predicted 115 OR gene models in the *P. fuscatus* genome. A combined *P. fuscatus* male and gyne (reproductive female) antennal transcriptome generated using Trinity (Haas et al. 2013) yielded 89 OR genes greater than 900 nucleotides in length. Some long Trinity genes contain multiple 7-transmembrane domains and likely represent concatenated OR genes. The small fraction of the *P. fuscatus* OR repertoire predicted by transcriptome assembly is consistent with previous observations that annotation of OR repertoires using only transcriptome data typically fails to recover all ORs (Karpe et al. 2016, 2017, 2021).

Manual gene annotation of *P. fuscatus* ORs recovered 231 gene models across 28 scaffolds (supplementary fig. S2, Supplementary Material online), of which 28 are pseudogenes and 10 are incomplete gene models (seven missing N termini, two missing C termini, and one missing both N and C termini). Since functional insect ORs are typically composed of

Table1. Summary of Odorant Receptor Gene Annotations in Five *Polistes* Genomes.

Species	Functional ORs[a]	Mean Length[b]	Mean TM TMHMM[c]	Mean TM Phobius[d]	OR Models	PSE	Partial Models
P. fuscatus	200	395 ± 15	5.95 ± 0.91	6.43 ± 1.13	231	28	10
P. metricus	204	396 ± 13	5.96 ± 0.85	6.45 ± 1.17	217	12	9
P. dorsalis	177	393 ± 20	5.90 ± 0.90	6.40 ± 1.21	203	16	24
P. canadensis	188	394 ± 17	5.95 ± 0.91	6.48 ± 1.20	235	13	59
P. dominula	180	392 ± 19	5.99 ± 0.88	6.59 ± 1.33	202	7	33

[a]OR gene models encoding proteins ≥ 300 amino acids in length.
[b]Mean length in amino acids (± SD).
[c]Mean transmembrane domains predicted by TMHMM.
[d]Mean transmembrane domains predicted by Phobius.

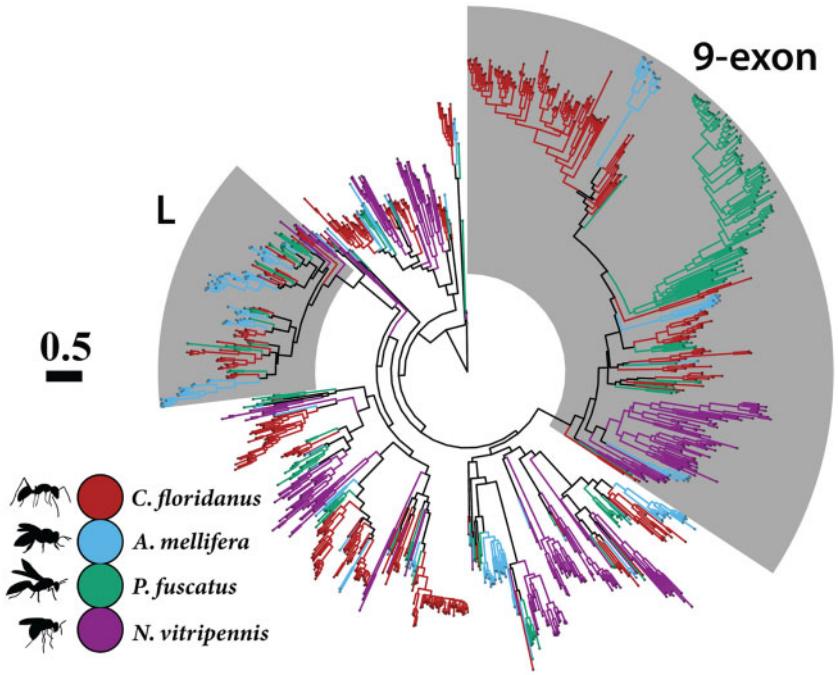

Fig. 3. Maximum likelihood OR protein tree constructed using data from four Hymenopterans (*Apis mellifera*, Robertson and Wanner 2006; *Camponotus floridanus*, Zhou et al. 2012; *Nasonia vitripennis*, Robertson et al. 2010). Branches are colored by species (Red: *C. floridanus*; Light blue: *A. mellifera*; Green: *P. fuscatus*; Purple: *N. vitripennis*). The L and 9-exon OR subfamilies are highlighted. Scale bar represents 0.5 mean substitutions per site.

400 amino acids, we defined gene models as putatively functional if they coded for proteins greater than or equal to 300 amino acids in length, even if the gene models were incomplete. In *P. fuscatus*, the 200 putatively functional gene models encode protein sequences with an average length of 395 ± 15 (SD) amino acids, and 198 of these gene models encode protein sequences greater than 350 amino acids in length (table 1). OR proteins possess seven transmembrane domains (Wicher 2015). The putatively functional *P. fuscatus* OR proteins possess on average 5.95 ± 0.91 (SD) transmembrane domains as predicted by TMHMM version 2.0c (Sonnhammer et al. 1998) and 6.43 ± 1.13 (SD) as predicted by Phobius version 1.01 (Käll et al. 2004). For comparison, transmembrane domain prediction in 61 *D. melanogaster* ORs coding for proteins greater than 375 amino acids in length found on average 5.77 ± 1.12 (SD) transmembrane domains as predicted by TMHMM version 2.0c and 6.18 ± 1.09 (SD) as predicted by Phobius version 1.01 (sequences from Supplemental Data 1, Supplementary Material online, in Hopf et al. 2015). The close match between the number of ORs predicted by neuroanatomy and the number recovered from manual annotation suggests that we have identified nearly all of the OR genes in *P. fuscatus*. The number of transmembrane domains predicted are comparable to annotations of *D. melanogaster* and approach the seven transmembrane domains expected for insect ORs. Manual OR gene annotation in *P. fuscatus* and four other *Polistes* genomes is summarized in table 1.

9-Exon OR Subfamily Expanded During the Evolution of Social Wasps

We conducted a Hymenoptera-wide analysis of OR evolution to test the prediction that the 9-exon OR subfamily was independently expanded during the evolution of eusociality in vespid wasps. By comparing the *P. fuscatus* OR repertoire to other Hymenopterans, our findings reinforce previous results showing that across Hymenopteran families, ORs evolve with

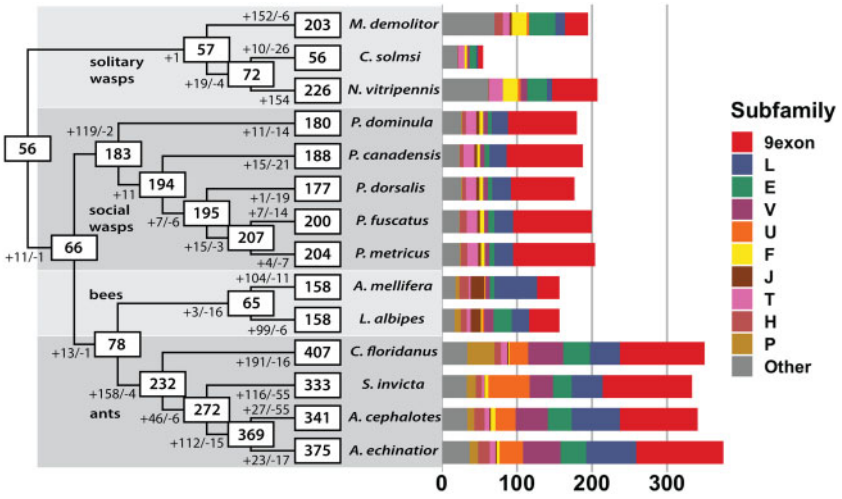

FIG. 4. Cladogram of Hymenoptera species showing estimated number of OR gene gain and loss events along branches and estimated size of ancestral and extant species OR repertoires in boxes. To the right is a bar chart showing numbers of ORs broken down by subfamily. Non-Polistine OR data are from Robertson et al. (2010) and Zhou et al. (2012, 2015). The set of intact ORs that were longer than 300 amino acids was used except for *C. floridanus* in the bar chart, where only ORs considered putatively functional by Zhou et al. (2012) were used.

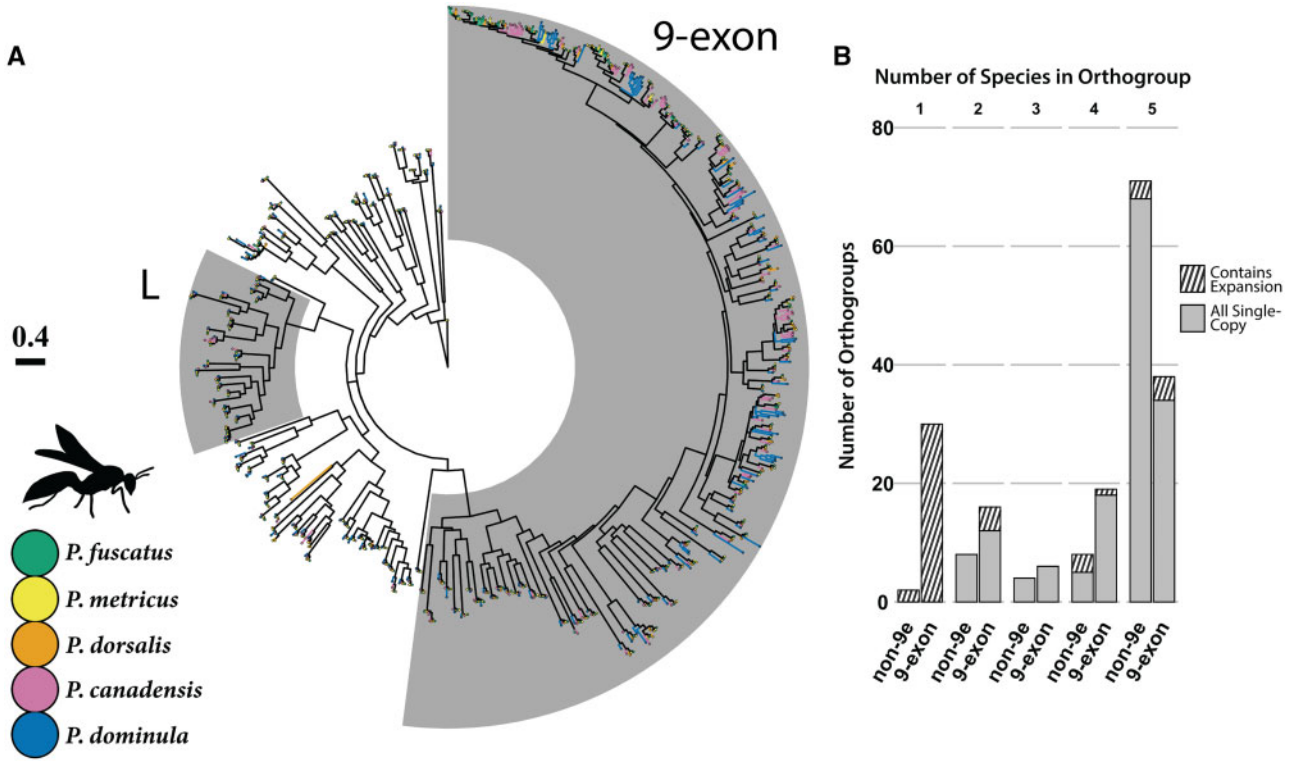

FIG. 5. (A) Maximum likelihood OR protein tree with branches colored by species (Green: *Polistes fuscatus*; Yellow: *P. metricus*; Orange: *P. dorsalis*; Magenta: *P. canadensis*; Blue: *P. dominula*). The L and 9-exon subfamilies are highlighted. Scale bar represents 0.4 mean substitutions per site. (B) Stacked bar chart showing the number of *Polistes* species (*x*-axis) represented in each orthologous group (*y*-axis), and whether or not each orthologous group is single copy (shaded bottom portion of bar) or contains an expansion in at least one species (top-striped portion of bar). Orthologous groups are split into two categories: non-9-exon orthologous groups ("non-9e" left bar) and 9-exon orthologous groups (right bar).

lineage-specific expansions of multiple OR subfamilies (fig. 3). Gene gain and loss events were predicted using NOTUNG (Chen et al. 2000) and mapped onto a species cladogram of 14 Hymenopterans (fig. 4). NOTUNG estimated an ancestral Apocritan repertoire of 56 ORs, which has expanded independently during the evolution of braconid wasps, ants, bees, and paper wasps (fig. 4). The 9-exon subfamily is commonly expanded across Hymenoptera (~90 genes on average), and comprises ~36% of social insect OR repertoires. The largest lineage-specific expansions of Hymenopteran 9-

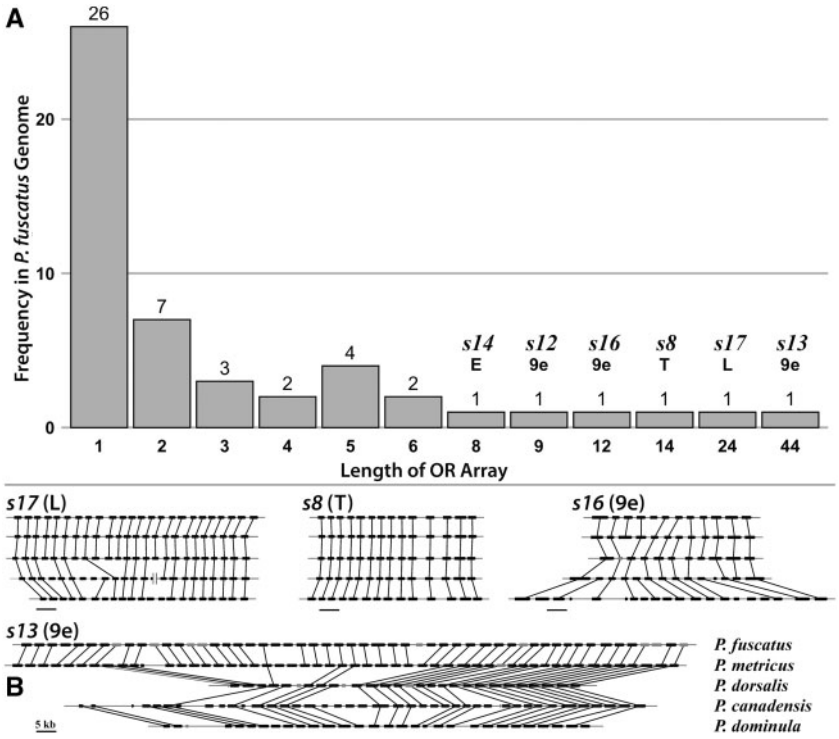

Fig. 6. (A) Frequency of OR gene singletons and tandem arrays in the *Polistes fuscatus* genome. 62% of ORs in *P. fuscatus* occur in tandem arrays of six or more genes. The longest tandem array is a 44 gene cluster on scaffold 13 (*s13*) containing 9-exon subfamily ORs. The first row of *x*-axis labels is the number of OR genes in a tandem array cluster, and the second row labels the OR subfamily and scaffold number (abbreviated *s#* in parentheses) of the six longest tandem arrays. (B) Genome alignments of four loci containing tandem arrays of OR genes in all *Polistes* species examined. Each alignment is labeled with the corresponding OR subfamily and *P. fuscatus* scaffold number. Black boxes represent putatively functional genes (≥300 amino acids) and gray boxes represent pseudogenes. Directionality of genes is denoted by curved corners at the 3′ (tail) end. Black lines connect orthologous genes between species. Genomic scaffolds are represented by horizontal, gray lines, and scaffold ends are represented by vertical gray lines. Scale bars beneath each alignment represent 5 kb.

exon ORs have occurred independently during the evolution of ants and social wasps. In *P. fuscatus*, this clade has expanded to 105 genes, comprising 53% of the OR gene set (fig. 4). Given the well-documented use of CHCs as signal molecules in *Polistes* (Singer 1998; Dani et al. 2001; Dani 2009; Beani et al. 2019), it is not surprising to find expansions in the putatively CHC-detecting 9-exon subfamily in this genus. Subfamilies L, T, H, E, and V have also expanded in *Polistes*, but not to the extent of the 9-exon OR subfamily.

The 9-Exon OR Subfamily Shows a Distinct Pattern of Orthology within *Polistes*

We next examined the evolutionary history of OR genes among the five *Polistes* species to reveal patterns of orthology and paralogy within subfamilies. Across the *Polistes* genus, most OR subfamilies are highly conserved (fig. 5*A*; supplementary fig. S3, Supplementary Material online). About 70% of non-9-exon family *P. fuscatus* ORs are in 1:1 orthology with all other *Polistes* species sampled as predicted by OrthoFinder (Emms and Kelly 2015) (supplementary table S3, Supplementary Material online). The remaining orthologous groups contain an expansion in one or more species (fig. 5*B*). Considering non-9-exon ORs, most ORs are shared by all five *Polistes* species examined, and most expansions are shared across all five species. Given that the species examined here span ~40 million years of divergence (Peters et al. 2017), the conservation of most of the OR repertoire is notable and may be related to the similarity of ecological and social niches found among *Polistes* wasps. While a common evolutionary history has led to large 9-exon OR complements in all *Polistes* species examined, lineage-specific gains and losses of 9-exon ORs account for most of the variation in OR repertoire size across *Polistes* species (fig. 4). In contrast with the other OR subfamilies, the 9-exon OR subfamily shows more lineage specificity with only 32% of *P. fuscatus* 9-exon ORs showing simple 1:1 orthology across all five *Polistes* examined (supplementary table S3, Supplementary Material online). Most 9-exon subfamily orthologous groups contain gene copies from four or fewer species, and lineage-specific expansions are more common in 9-exon OR orthologous groups (fig. 5*B*). The relative lack of orthology among 9-exon OR genes compared to the rest of the OR gene subfamilies suggests unique evolutionary processes shaping 9-exon ORs.

Microsynteny Reveals Recent Birth and Death Events in *Polistes* 9-Exon OR Subfamily

Expanded gene families often occur as tandem arrays, a genomic architecture that can contribute to increased rates of gene birth and death, increasing copy number variation among species (Ohno 1970). Therefore, we examined how

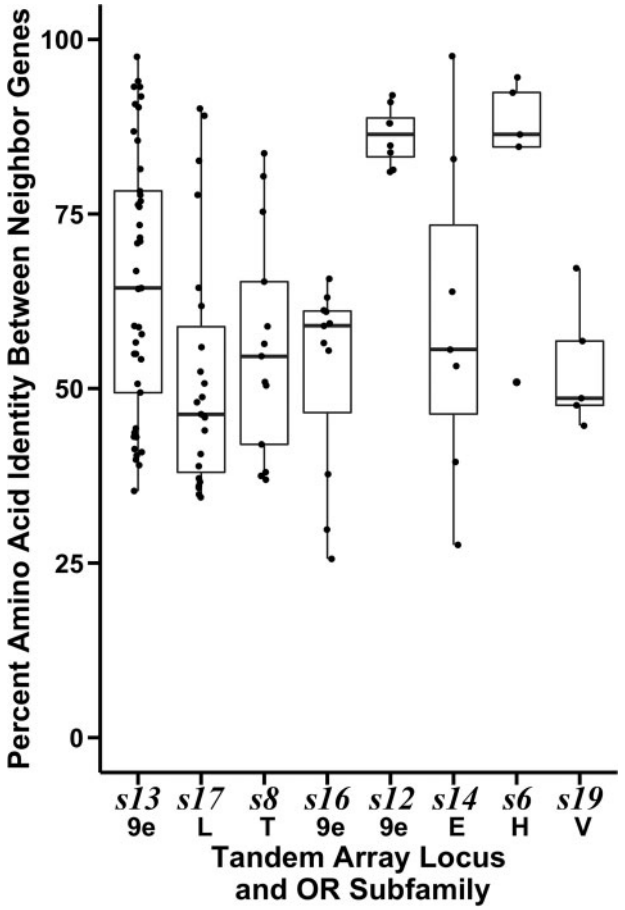

FIG. 7. Percent amino acid identity between neighboring genes at eight loci containing the longest OR gene tandem arrays in the *P. fuscatus* genome. Arrays are ordered by length in gene number, from longest (44 9-exon subfamily ORs in the *s13* tandem array) to shortest (6 H subfamily ORs in the *s6* tandem array and 6 V subfamily ORs in the *s19* tandem array).

genomic organization varies between OR subfamilies in *Polistes* species to generate insights into the molecular evolutionary mechanisms shaping OR subfamily function. Genomic organization of ORs across *Polistes* is consistent with a model of birth and death evolution shaping OR repertoires. As in bees, gene gain and loss at a small number of loci containing tandem arrays are responsible for most copy number variation in the OR family across closely related species (Brand and Ramírez 2017). In *P. fuscatus*, 62% of ORs occur in tandem arrays of six or more genes (fig. 6*A*). The frequency of tandem arrays and the tail-to-head orientations of neighboring genes point to tandem duplication as the primary mechanism of OR expansion, likely caused by non-allelic homologous recombination (Lynch 2007; Ramdya and Benton 2010). We examined microsynteny of OR genes and pseudogenes in the four longest tandem arrays of ORs in *Polistes*. Gene birth and death events have resulted in more complex orthology among genes in orthologous 9-exon OR arrays compared to tandem arrays of L and T subfamily ORs in *Polistes* genomes (fig. 6*B*). The longest OR gene tandem array in *P. fuscatus* is comprised of 44 genes in the 9-exon subfamily on scaffold 13 (*s13*), which corresponds to homologous arrays of 50 genes in *P. metricus*, 25 genes in *P. dorsalis*, 33 genes in *P. canadensis*, and 29 genes in *P. dominula*. Only 34% of *P. fuscatus* ORs in this array have orthologs across all *Polistes* species sampled, and collinear orthologs in this array are frequently interrupted by inparalogs (fig. 6*B*). The second longest OR gene tandem array in *P. fuscatus* contains 24 ORs in the L subfamily on scaffold 17 (*s17*), and these ORs show 1:1 orthology across *P. fuscatus*, *P. metricus*, and *P. dorsalis*, while *P. canadensis* possesses an array of ~23 genes split between two scaffolds, and *P. dominula* possesses an array of 21 ORs at this locus. This tandem array, widely expanded across Hymenoptera, has been expanded and conserved across *Polistes*. The T subfamily, located on scaffold 8 (*s8*) of the *P. fuscatus* genome, is composed of 14 tandemly arrayed genes that show 1:1 orthology across five *Polistes* (fig. 6*B*). Differences between OR subfamilies in patterns of orthology within microsyntenic regions highlight the unique evolutionary processes shaping 9-exon OR evolution in paper wasps. At the same time, the orthology of collinear genes within syntenic L and T subfamily tandem arrays across all examined *Polistes* species highlights the strong conservation of much of the *Polistes* OR repertoire.

Microsynteny analysis suggests a process of ongoing gene turnover in 9-exon arrays but stasis in most other expanded subfamilies. More recent turnover should be associated with higher pairwise amino acid identity between neighboring genes in an array if they are the result of recent duplication events (Ohno 1970; Bohbot et al. 2007; Engsontia et al. 2015). To explore the relationship between amino acid divergence and tandem array locus, we compared the mean percent amino acid identity among neighboring genes within an array between the eight loci containing the longest tandem arrays of ORs in the *P. fuscatus* genome using one-way ANOVA (fig. 7). Mean percent amino acid identity of neighboring genes was significantly separated by OR array identity (DF = 7; $F = 5.39$; *P*-value = 2.67e−05). Differences between particular OR tandem arrays were identified using Tukey HSD post hoc tests. The mean percent amino acid identity among neighboring genes within one tandem array of nine 9-exon ORs on scaffold 12 (*s12*) of the *P. fuscatus* genome is higher than in the *s13* 9-exon array (*P* Adj = 0.04586), the *s17* L array (*P* Adj = 0.00013), the *s8* T array (*P* Adj = 0.00458), the *s16* 9-exon array (*P* Adj = 0.00124), and the *s19* V array (*P* Adj = 0.02247). The *s12* 9-exon OR array is composed of a larger proportion of pseudogenes (5 PSE, 10 intact gene models) than the other two 9-exon arrays (*s13*: 9 PSE, 44 intact gene models; *s16*: 0 PSE, 12 intact gene models). ORs in the *P. fuscatus s12* 9-exon array lack clear orthologous relationships with ORs in species other than *P. metricus*. Taken together, the high within array sequence similarity, high frequency of pseudogenes, and low orthology exhibited by this array indicate that it is the result of one or more recent gene duplication events since the divergence of *P. fuscatus* and *P. metricus* from the other three *Polistes* species. The *s6* H subfamily array also shows higher amino acid sequence identity among neighboring genes than the *s17* L subfamily array (*P*-value Adj = 0.01625) and the *s16* 9-exon array (*P*-value Adj = 0.04038). Increased amino acid similarity may also occur

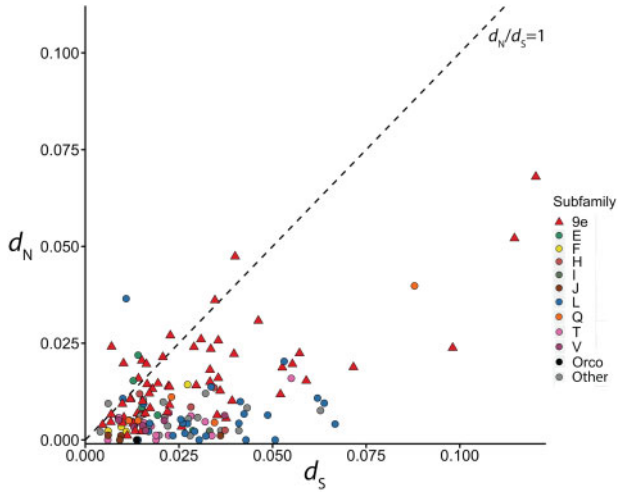

FIG. 8. The values of d_S (*x*-axis) and d_N (*y*-axis) from pairwise alignments of *Polistes fuscatus* and *P. dorsalis* 1:1 orthologs. Values of d_N are elevated in the 9-exon OR subfamily (data points represented by red triangles) relative to other OR subfamilies (data points represented by circles). The diagonal line represents a line of equality with slope of 1.

within older tandem arrays as a result of gene conversion (Nagawa et al. 2002). However, we searched for gene conversion using GENECONV (Sawyer 1989) and did not detect gene conversion events within the *s12* 9-exon array or in the *s6* H array after Bonferroni correction. Patterns of genomic organization of OR genes in *Polistes* genomes lead to the conclusion that gene gain and loss in the 9-exon OR subfamily is an ongoing process within this genus, in contrast to the stable and conserved tandem arrays in most other OR subfamilies.

Positive Selection in Expanded OR Subfamilies and Accelerated Evolution of 9-Exon ORs

We examined the variation in omega (d_N/d_S) and used codon models to characterize sequence evolution of *Polistes* OR genes. We were especially interested in patterns of selection in the 9-exon OR clade. Consecutive analyses of *Polistes* OR subfamilies using HyPhy adaptive branch-site random effects likelihood (aBSREL) model (Smith et al. 2015; Pond et al. 2020) detected eight branches under episodic positive selection, all in OR subfamilies with expansions: three branches in the 9-exon subfamily (0.33% of 918 9-exon subfamily branches; supplementary fig. S4, Supplementary Material online); three branches in the L subfamily (1.28% of 234 L subfamily branches; supplementary fig. S5, Supplementary Material online); one branch in the E subfamily (1.67% of 60 E subfamily branches; supplementary fig. S6, Supplementary Material online); and one branch in the H subfamily (1.54% of 65 H subfamily branches; supplementary fig. S7, Supplementary Material online). This supports the hypothesis that gene duplication releases duplicate genes from selective constraints, allowing duplicate sequences to evolve towards other evolutionary optima (Ohno 1970; Saad et al. 2018).

To visualize the range of patterns of synonymous and nonsynonymous substitutions in *Polistes* ORs, we computed the values of d_N and d_S for pairwise alignments of 151 single copy orthologs between *P. fuscatus* and *P. dorsalis* (fig. 8) using model yn00 of PAML (Yang 2007). Values of d_N are significantly higher in 9-exon (mean $d_N = 0.015$) compared to other OR ortholog pairs (mean $d_N = 0.006$) (Welch Two Sample *t*-test; *P*-value = 5.317e−07). Values of d_S are not significantly elevated among 9-exon ortholog pairs compared to other OR subfamilies (mean $d_S = 0.029$ in 9-exon ORs and 0.025 in non-9-exon ORs; *P*-value = 0.343). Omega values (d_N/d_S) greater than 1 are often considered evidence of positive selection, while $d_N/d_S = 1$ corresponds to neutral drift, and $d_N/d_S < 1$ is evidence of negative (purifying) selection. The omega value (d_N/d_S) for the majority of OR genes is less than one, suggesting negative selection (mean omega = 0.454). However, omega is significantly higher in 9-exon ORs than in non-9-exon ORs, indicating that negative selection is weaker on 9-exon ORs (Welch Two Sample *t*-test: mean omega = 0.644 in 9-exon ORs ($N = 62$) and 0.32 in non-9-exon ORs ($N = 89$); *P*-value = 8.027e−05). This 1:1 orthology analysis excludes patterns of molecular evolution among genes with more complex orthology relationships, though an analysis considering all orthogroups identified by OrthoFinder containing at least four genes ($N = 145$) confirms the pattern. Estimates of omega by M0 in CodeML are significantly higher in 9-exon orthogroups than in non-9-exon orthogroups (Welch Two Sample *t*-test: mean omega = 0.407 in 9-exon ORs ($N = 66$) and 0.189 in non-9-exon ORs ($N = 79$); *P*-value < 2.2e−16) (supplementary fig. S8, Supplementary Material online). An elevated omega ratio in 9-exon 1:1 ortholog pairs and orthogroups implies that either a relaxation of negative selection or an intensification of positive selection is responsible for sequence evolution in the 9-exon relative to other OR subfamilies. We explicitly tested the hypothesis that relaxed negative selection is responsible for higher omega values in branches of the 9-exon OR clade compared to branches in other OR subfamilies using HyPhy RELAX (Wertheim et al. 2015; Pond et al. 2020). This analysis found a significant pattern of reduced selection intensity in the 9-exon OR clade compared to the rest of the *Polistes* OR tree (LRT = 505.81; mean selection intensity parameter $k = 0.48$; *P*-value < 0.0001). Relaxed selection in the 9-exon OR subfamily may allow ORs to explore phenotypic space and develop novel response spectra for behaviorally relevant chemical signals.

Discussion

Expansion of 9-Exon OR Subfamily during Independent Evolution of Sociality in Wasps

By carefully annotating the OR repertoires of five social wasp species spanning ~40 million years of divergence in the *Polistes* genus, this study adds a higher resolution lens to our view of the evolution of social insect ORs. During the diversification of *Polistes*, evolutionary patterns show genus-wide conservation of their ~200 ORs except for the 9-exon genes, which show elevated turnover and lower sequence conservation. The 9-exon OR subfamily has expanded in paper wasps, and now makes up over half of the *Polistes* OR

gene set. Social and ecological niches are relatively conserved within *Polistes*, though there is considerable variation in social behavior and ecological niches among vespid wasps (Ross and Matthews 1991; O'Neill 2001). An analysis of three hornet genomes suggested that the highly eusocial hornets may have even larger OR repertoires compared to the primitively eusocial *Polistes* (Harrop et al. 2020). That analysis recovered less than half of the ORs reported here for *Polistes*, likely due to a lack of manual annotation informed by antennal transcriptome data, suggesting that hornets may have larger OR repertoires than reported. Evidence from the hornet *Vespa velutina*, including the discovery of ~265 antennal lobe glomeruli, indicates that the hornet OR repertoire has expanded (Couto et al. 2016). Interestingly, 96 glomeruli populate the T_B cluster in *V. velutina*, an antennal lobe region innervated by CHC-detecting sensilla basiconica and proposed to be homologous with the T6 antennal lobe cluster of ants (Couto et al. 2017). Given that the number of T6 glomeruli correlates with the number of 9-exon OR genes expressed in ant antennae, the large number of T_B glomeruli in hornets strongly suggests an expanded complement of 9-exon ORs (McKenzie et al. 2016). Future analysis of additional genomes and antennal transcriptomes of diverse social and solitary vespid wasps will allow further examination of the relationship between social behavior and OR subfamily expansion.

Combinatorial Coding of CHCs by 9-Exon ORs Facilitates Recognition

Electrophysiological deorphanization studies of 9-exon ORs in the ant *Harpegnathos saltator* offer key insights into how 9-exon OR coding might relate to gene expansion. Through combinatorial coding, 9-exon ORs can detect a large variety of structurally diverse CHCs. Pask et al. (2017) examined 22 *H. saltator* 9-exon ORs, a subset of the 118 annotated 9-exon ORs in this species, and found that 9-exon ORs were responsive to CHCs, and overlapped in their responses to multiple CHC compounds. The combined responses of these 22 ORs to CHC extracts from different castes were sufficient to map the CHC profiles of males, workers, and reproductive females (gamergates) to separate regions of a 22-dimensional receptor space (Pask et al. 2017). This highlights the ability of 9-exon ORs to facilitate social recognition by combinatorial coding. In social insect colonies, CHC variation holds information at multiple levels of conspecific recognition, from inter-colony nestmate recognition to within colony individual recognition (Greene and Gordon 2003; d'Ettorre and Heinze 2005; d'Ettorre and Moore 2008; Leonhardt et al. 2016). Expansion of the 9-exon OR subfamily might result from selection for more combinations of ORs that together can discriminate between subtle qualitative and quantitative variations in CHC blends of conspecifics. Nest-specific quantitative variation in CHCs has been documented across *Polistes* species (Espelie et al. 1990; Singer et al. 1992; Espelie et al. 1994; Layton et al. 1994), but the molecular mechanisms underlying nestmate recognition in *Polistes* are still obscure. Increased copy number of 9-exon ORs may not only expand the qualitative range of compounds perceived by paper wasps, but also the perceived quantitative olfactory space, since wasps may be able to discern unique concentration differences between CHC blends as a result of the combined action of 9-exon ORs with various response thresholds. Gene duplication can also promote regulatory diversification (Kucharski et al. 2016; Dyson and Goodisman 2020). In *P. metricus*, CHCs vary between castes and across stages of the colony cycle (Toth et al. 2014). Regulatory subfunctionalization of duplicate ORs could be responsible for caste- and colony phase-specific expression of ORs involved in detecting caste-specific and seasonally variable CHCs. In addition to adaptive expansion of ORs, neutral processes contribute to OR gene birth-and-death events. There may be an advantage for a large 9-exon OR gene copy number up to a point, followed by random gene duplication and deletion around this optimal copy number. This random genomic drift has been proposed to shape mammalian olfactory receptor evolution and copy number variation in other large multigene families (Nei 2007; but see Hayden et al. 2010). Indeed, we find evidence of relaxed selection on the 9-exon OR subfamily compared to other wasp OR subfamilies, which is consistent with predictions for the evolution of combinatorial coding (Andersson et al. 2015). Mutations that slightly alter the response profiles of functionally redundant ORs may not be eliminated by negative selection, since other ORs can help compensate (Fishilevich et al. 2005; Keller and Vosshall 2007).

Evolution of ORs Reflects Distinct Chemosensory Ecologies of Species

Social insect species differ in their level of sociality and extent of olfactory recognition abilities (Stuart 1988; Page et al. 1991; d'Ettorre and Moore 2008; Peeters and Liebig 2009; Rehan and Toth 2015). Some aspects of the *Polistes* colony cycle vary across species. For example, the average number of cooperative foundresses varies from 1 to ~6, and average sizes of mature nests may vary from ~60 cells in *P. metricus* to ~490 cells in *P. annularis* (Rabb 1960; Downing and Jeanne 1986; Reeve 1991; Sheehan et al. 2015; Miller SE et al. 2018). Increased 9-exon OR copy number may facilitate complex olfactory recognition in species with larger colony sizes, higher cooperative nest-founding rates, and greater sympatry with related species. However, expansions of 9-exon ORs are not exclusive to social wasps, suggesting that the specific chemical ecology of an insect is a more influential factor shaping OR evolution than level of sociality (Karpe et al. 2017). Furthermore, a meta-analysis found that the complexity of CHC phenotypes does not differ between social and solitary Hymenopteran species (Kather and Martin 2015). The CHC profile of *Nasonia vitripennis* includes at least 52 CHC compounds, and detection of CHCs on prey items may help *Microplitis* identify prey (Lewis et al. 1988; Niehuis et al. 2011). The need for parasitoid wasps to perceive CHCs could explain why genomes of *N. vitripennis* and *M. demolitor* exhibit expansions in the 9-exon OR subfamily. The OR repertoire of the fig wasp *Ceratosolen solmsi* is about one-third of the size of those described in *N. vitripennis* and *M. demolitor*, which likely reflects the specialized sensory demands of identifying one host plant species (Xiao et al. 2013). In general, the 9-exon OR subfamily comprises a smaller proportion of the

OR gene set in bees, which do not predate insects, relative to ants and wasps. The extreme expansion of 9-exon ORs in the myrmecophagous clonal raider ant relative to other ant species provides further evidence suggesting that the need to detect insect prey may influence 9-exon OR gene content.

Dietary selective pressures may play a part in shaping the evolution of *Polistes* ORs. Paper wasps appear to primarily predate Lepidoptera larvae, and sympatric *Polistes* species predate overlapping sets of prey species (Rabb and Lawson 1957; Rabb 1960; Southon et al. 2019). *P. dominula* workers were found to predate a wider range of insects than other *Polistes* species (Cervo et al. 2000). Variation in the OR repertoire could underlie variation in foraging behavior between *Polistes* species.

Lineage-Specific Molecular Evolution of *Polistes* 9-Exon ORs

Most expanded OR subfamilies are highly conserved in copy number across five *Polistes* species, with the exception of the 9-exon OR subfamily. In particular, one portion of the 9-exon subfamily arranged in a tandem array (*P. fuscatus* 9e *s13*; fig. 6*B*) has experienced dynamic evolution. What selective pressures might drive rapid gain and loss of 9-exon ORs? Divergent chemical signaling between species may lead to gene turnover as 9-exon OR evolution tracks evolutionarily labile chemical signals. For example, *P. fuscatus* and *P. metricus* are relatively closely related, and both species possess CHC profiles consisting of linear and methyl-branched alkanes (Espelie et al. 1990; 1994). However, the *P. fuscatus* CHC profile includes a higher proportion of alkenes than *P. metricus* or *P. dominulus*, and the position of the methylated carbon of methyl-branched alkanes is sometimes shifted between species (Espelie et al. 1990; Singer et al. 1992; Espelie et al. 1994; Layton et al. 1994). Ant 9-exon ORs respond differently to subtle variations in CHC structure (Pask et al. 2017). Between closely related *Polistes* species, structural isomers of methyl-branched alkanes probably activate different ensembles of ORs.

If a chemical evolves new behavioral relevance in a population, gene duplication could allow the olfactory system to explore chemical space in the direction of this compound. HyPhy aBSREL analyses identified eight branches in expanded OR subfamilies, including the 9-exon subfamily, that have undergone positive selection during the last ~40 million years, consistent with neofunctionalization or subfunctionalization of duplicated genes. Signatures of positive selection on OR genes may indicate directional selection to perceive species-specific chemical signals. Perception of species-specific CHCs might be important in mate compatibility recognition. In *Polistes*, mating occurs at sites defended by males and visited by females of multiple species (Post and Jeanne 1984, 2010; Reed and Landolt 1990). However, the frequency of interspecific mating is low, suggesting *Polistes* use vision and/or olfaction to inform their mating decisions (Miller et al. 2019). Duplication and deletion of ORs would facilitate evolution of species-specific chemical signaling systems that could contribute to reproductive isolation of sympatric species. If a chemical signal is lost in a species, the corresponding ORs may become obsolete, and would be expected to pseudogenize and be purged from the genome. Duplication and deletion of ORs could also lead to species-specific chemical signaling in the absence of evolutionary change in chemical signals (Cande et al. 2013). However, OR evolution is not strictly necessary for such a difference to evolve between species, and circuit-level changes can prescribe new valence to chemical signals that are shared between species and perceived by common peripheral receptors (Seeholzer et al. 2018).

Conservation of Most OR Subfamilies Suggests Conserved Functions

Aside from the 9-exon OR subfamily, gene expansions have occurred in subfamilies L, T, H, E, and V (fig. 4). A larger variety of ORs relaying information through odorant receptor neurons (ORNs) to a larger number of antennal lobe glomeruli will increase sensory acuity in any olfactory discrimination task, social, or otherwise. An ancient locus of tandemly duplicated L subfamily ORs observed across social insects has expanded in *Polistes*, although to a lesser extent than in other social insects (~50 L subfamily ORs in honeybee and ants, 25 L subfamily ORs in a tandem array on *P. fuscatus* scaffold 17). ORs in the L subfamily are thought to detect queen pheromone components and fatty acids in bees as well as CHCs in ants (Wanner et al. 2007; Karpe et al. 2016; Slone et al. 2017). The T subfamily has expanded to a greater degree in *P. fuscatus* (14 genes) than in ants (~7 genes) and the honeybee (2 genes), but no ORs in this clade have been functionally characterized. The *P. fuscatus* genome encodes nine H subfamily ORs, which are putative floral odorant detectors in bees, and which also respond to CHCs and other general odorants in ants (Claudianos et al. 2014; Slone et al. 2017). Fatty acids and volatile organic compounds are produced by flowers that wasps rely on as a source of carbohydrates (Raguso 2008). Expansions in several OR subfamilies may increase olfactory discrimination of chemicals with diverse behavioral relevance. *Polistes* species are distributed globally in temperate and tropical regions, occupying similar social and ecological niches as generalist predators of Lepidoptera and floral foragers that form primitively eusocial societies (Reeve 1991; Richter 2000). A conserved subset of the OR repertoire may perform common functions in conserved behaviors across paper wasp species. High levels of OR conservation are also consistent with a specialist molecular function of an OR in a dedicated channel of olfaction (Andersson et al. 2015).

Distinct Patterns of OR Evolution within the Same Genome

In paper wasps, we report both highly conserved OR expansions similar to those seen in *Drosophila* as well as elevated gene turnover and drift among the 9-exon ORs, reminiscent of a more mammal-like evolutionary pattern. The differences between the conserved OR repertoires in *Drosophila* and the more dynamic evolution of mammal OR gene families have given rise to speculation about the relationship between OR function and evolution (Nozawa and Nei 2007; Andersson

et al. 2015). If the highly dynamic clades of 9-exon ORs of social wasps are involved in more combinatorial coding compared to other more conserved 9-exon or non-9-exon ORs, that would indicate a link between molecular evolution of ORs and neural coding. Further investigations into the relative tuning of 9-exon as well as more conserved ORs in social wasps and other social insects provide a promising research direction to investigate the links between molecular evolutionary patterns, OR tuning, and neural coding.

Materials and Methods

Antennal Lobe Imaging

Antennal lobe glomeruli of male and female *P. fuscatus* wasps were stained with anti-synapsin and imaged using a confocal laser scanning microscope. Details of the immunocytochemistry and imaging are included as Supplementary Material online.

Gene Annotation

The genomes and annotations of *P. canadensis*, *P. dominula*, *P. fuscatus*, *P. dorsalis*, and *P. metricus* were accessed through NCBI (Patalano et al. 2015; Standage et al. 2016; Miller SE et al. 2020). Coding regions of ORs were identified by using TBLASTN (Altschul et al. 1997) with a sample of OR proteins from 19 insect species used as query sequences: *Atta cephalotes*, *Acromyrmex echinatior*, *Apis mellifera*, *Camponotus floridanus*, *Cardiocondyla obscurior*, *Ceratosolen solmsi*, *Drosophila melanogaster*, *Eulaema bombiformis*, *Euglossa dilemma*, *Euglossa flammea*, *Euglossa imperialis*, *Eulaema meriana*, *Eufriesea mexicana*, *Lasioglossum albipes*, *Microplitis demolitor*, *Monomorium pharaonis*, *Melipona quadrifasciata*, *Nasonia vitripennis*, *Solenopsis invicta* (Robertson et al. 2003, 2010; Zhou et al. 2012, 2015; Brand and Ramírez 2017; McKenzie and Kronauer 2018). Genomes were queried iteratively with TBLASTN, adding newly annotated *Polistes* ORs to the query file, until no new OR coding regions were identified. To guide annotation of exon–intron boundaries, antennal mRNA from *P. fuscatus* males and females (gynes) was mapped to *P. fuscatus*, *P. metricus*, and *P. dorsalis* genomes using STAR (Dobin et al. 2013) and assembled into transcripts using Trinity (Haas et al. 2013) (supplementary table S2, Supplementary Material online). Predicted transcripts were aligned to genomes using BLAT. Uncertain gene models in *P. metricus*, *P. dorsalis*, *P. canadensis*,and *P. dominula* were aligned to their orthologs in *P. fuscatus* using Muscle version 3.8.425 with maximum four iterations (Edgar 2004), and gene models were manually adjusted. All annotation evidence was imported into Geneious v11.1.5 genome browser for manual annotation. The majority of apparently functional ORs that were not detected by the automated annotation and required extensive manual curation were 9-exon subfamily receptors (e.g. supplementary fig. S9, Supplementary Material online). Gene models were called pseudogenes if they exhibited frame-shift mutations, premature stop codons, or unacceptable 5′ donor or 3′ receptor splice sites. Transmembrane helices of all putatively functional ORs (>300 amino acids) were predicted using TMHMM version 2.0c (Sonnhammer et al. 1998) and Phobius version 1.01 (Käll et al. 2004). Throughout the main text, "putatively functional ORs" are OR proteins at least 300 amino acids in length. Details of mRNA library preparation, sequencing, read mapping, and manual gene annotation are included in Supplementary Material online.

Phylogenetic Reconstruction

Phylogenetic trees were constructed using RAxML (Jones et al. 1992; Stamatakis 2014). Gene duplication and loss events were reconstructed by reconciling a gene tree with a species tree in NOTUNG version 2.9.1.3 (Durand et al. 2006; Vernot et al. 2008). Orthologous genes were determined using OrthoFinder (Emms and Kelly 2015), bootstrap support, and microsynteny. Phylogenetic reconstruction methods are explained in detail in Supplementary Material online.

Genomic Organization

OR genes and pseudogenes were considered to be in a tandem array if they were uninterrupted by non-OR genes and were within 5 kb of each other. The lengths of OR arrays correspond to the number of putatively functional ORs and exclude the pseudogenes contained within the array. The pairwise percent amino acid identity between neighbors in an array was calculated using only putatively functional ORs that neighbored another putatively functional OR within 5 kb. Details of analyses of genomic organization are included in the Supplementary Material online.

Sequence Analyses

All putatively functional *Polistes* ORs greater than 350 amino acids in length were used in analyses of selection. The adaptive branch-site relative effects likelihood model (aBSREL) was used to test for signatures of episodic diversifying positive selection in HyPhy version 2.5.15 (Smith et al. 2015; Pond et al. 2020). Values of pairwise d_N/d_S for orthologs shared by *P. fuscatus* and *P. dorsalis* were estimated using PAML version 4.9 (Yang 2007) program yn00 with the Yang and Nielsen (2000) method. Values of d_N/d_S for *Polistes* orthogroups were estimated using PAML version 4.9 program CodeML with the M0 (one-ratio) model (Yang 2007). Finally, RELAX was run in HyPhy version 2.5.15 to test for relaxed negative (relaxed purifying) selection (Wertheim et al. 2015; Pond et al. 2020). RELAX was used to test for relaxed negative selection on 9-exon OR branches relative to non-9-exon OR branches in the *Polistes* OR tree. Gene subfamily codon alignments used in the above aBSREL analysis were tested with GENECONV to identify gene conversion (Sawyer 1989). Details of sequence analyses are included in the Supplementary Material online.

Data Availability

The genome assemblies analyzed in this article are available on Genbank (see supplementary table S1, Supplementary Material online). Gene models, amino acid sequences, and nucleotide sequences underlying this article, as well as alignments analyzed in HyPhy aBSREL and RELAX selection

analyses and phylogenetic trees in newick format, are available in its Supplementary Material online.

Acknowledgments

We thank Qi Sun and the Computational Biology Service Unit of the Cornell Life Sciences Core Laboratories Center for making software available on BioHPC and for providing helpful advice regarding gene annotation. We thank Brook Luers and the Cornell Statistical Consulting Unit, who provided statistics consultation. We thank Matt Einhorn and Judah Sklan for help troubleshooting scripts and Caitlin Miller for feedback on the manuscript. Finally, we thank two anonymous reviewers for critically reading and improving the manuscript.

Funding

This work was supported by National Science Foundation [Graduate Research Fellowship Program grant number DGE-1650441 to A.W.L., CAREER grant number DEB-1750394 to M.J.S.], National Institutes of Health [grant numbers DP2-GM128202 to M.J.S., S10OD018516 to the Cornell University Biotechnology Resource Center], and New York State Stem Cell Science [grant number CO29155 to Cornell BRC].

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25(17):3389–3402.

Andersson MN, Löfstedt C, Newcomb RD. 2015. Insect olfaction and the evolution of receptor tuning. *Front Ecol Evol.* 3(53):1–14.

Arnold G, Masson C, Budharugsa S. 1985. Comparative study of the antennal lobes and their afferent pathway in the worker bee and the drone (*Apis mellifera*). *Cell Tissue Res.* 242(3):593–605.

Beani L, Bagnères A-G, Elia M, Petrocelli I, Cappa F, Lorenzi MC. 2019. Cuticular hydrocarbons as cues of sex and health condition in *Polistes dominula* wasps. *Insect Soc.* 66(4):543–553.

Benton R. 2015. Multigene family evolution: perspectives from insect chemoreceptors. *Trends Ecol Evol.* 30(10):590–600.

Blomquist GJ, Bagnères A-G. 2010. Insect hydrocarbons. Cambridge: Cambridge University Press.

Bohbot J, Pitts RJ, Kwon H-W, Rützler M, Robertson HM, Zwiebel LJ. 2007. Molecular characterization of the *Aedes aegypti* receptor gene family. *Insect Mol Biol.* 16(5):525–537.

Brand P, Ramírez SR. 2017. The evolutionary dynamics of the odorant receptor gene family in corbiculate bees. *Genome Biol Evol.* 9(8):2023–2036.

Brand P, Ramírez SR, Leese F, Quezada-Euan JG, Tollrian R, 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):176.

Brand P, Robertson HM, Lin W, Pothula R, Klingeman WE, Jurat-Fuentes JL, Johnson BR. 2018. The origin of the odorant receptor gene family in insects. *eLife.* 7:e38340.

Butterwick JA, del Mármol J, Kim KH, Kahlson MA, Rogow JA, Walz T, Ruta V. 2018. Cryo-EM structure of the insect olfactory receptor Orco. *Nature.* 560(7719):447–467.

Cande J, Prud'homme B, Gompel N. 2013. Smells like evolution: the role of chemoreceptor evolution in behavioral change. *Curr Opin Neurobiol.* 23(1):152–158.

Carey AF, Wang G, Su C-Y, Zwiebel LJ, Carlson JR. 2010. Odorant reception in the malaria mosquito *Anopheles gambiae*. *Nature.* 464(7285):66–71.

Cervo R, Zacchi F, Turillazzi S. 2000. *Polistes dominulus* (Hymenoptera, Vespidae) invading North America: some hypotheses for its rapid spread. *Insectes Soc.* 47:155–157.

Chen K, Durand D, Farach-Colton M. 2000. NOTUNG: a program for dating gene duplications and optimizing gene family trees. *J Comput Biol.* 7(3-4):429–447.

Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman TC, Kellis M, Gelbart W, Iyer VN, Drosophila 12 Genomes Consortium, et al. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature.* 450(7167):203–218.

Claudianos C, Lim J, Young M, Yan S, Cristino AS, Newcomb RD, Gunasekaran N, Reinhard J. 2014. Odor memories regulate olfactory receptor expression in the sensory periphery. *Eur J Neurosci.* 39(10):1642–1654.

Cohanim AB, Amsalem E, Saad R, Shoemaker D, Privman E. 2018. Evolution of olfactory functions on the fire ant social chromosome. *Genome Biol Evol.* 10(11):2947–2960.

Couto A, Alenius M, Dickson BJ. 2005. Molecular, anatomical, and functional organization of the *Drosophila* olfactory system. *Curr Biol.* 15(17):1535–1547.

Couto A, Lapeyre B, Thiéry D, Sandoz J-C. 2016. Olfactory pathway of the hornet *Vespa velutina*: new insights into the evolution of the Hymenopteran antennal lobe. *J Comp Neurol.* 524(11):2335–2359.

Couto A, Mitra A, Thiéry D, Marion-Poll F, Sandoz J-C. 2017. Hornets have it: a conserved olfactory subsystem for social recognition in Hymenoptera? *Front Neuroanat.* 11:48.

d'Ettorre P, Moore AJ. 2008. Sociobiology of communication: an interdisciplinary perspective. In:d'Ettorre P, Hughes DP, editors. Chapter 5: chemical communication and the coordination of social interactions in insects. New York: Oxford University Press. p. 81–96.

Dani FR. 2009. Cuticular lipids as semiochemicals in paper wasps and other social insects. *Ann Zool Fennici.* 43:500–514.

Dani FR, Jones GR, Destri S, Spence SH, Turillazzi S. 2001. Deciphering the recognition signature within the cuticular chemical profile of paper wasps. *AnimBehav.* 62(1):165–171.

Dani FR, Turillazzi S. 2018. Chemical communication and reproduction partitioning in social wasps. *J Chem Ecol.* 44(9):796–804.

Dapporto L, Santini A, Dani FR, Turillazzi S. 2007. Workers of a *Polistes* paper wasp detect the presence of their queen by chemical cues. *Chem Senses.* 32(8):795–802.

D'Ettorre P, Heinze J. 2005. Individual recognition in ant queens. *Curr Biol.* 15(23):2170–2174.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 29(1):15–21.

Downing HA, Jeanne RL. 1986. Intra- and interspecific variation in nest architecture in the paper wasp *Polistes* (Hymenoptera, Vespidae). *Ins Soc.* 33(4):422–443.

Durand D, Halldórsson BV, Vernot B. 2006. A Hybrid Micro–Macroevolutionary Approach to Gene Tree Reconstruction. *J Comput Biol.* 13(2):320–335.

Dweck HKM, Ebrahim SAM, Farhan A, Hansson BS, Stensmyr MC. 2015. Olfactory proxy detection of dietary antioxidants in *Drosophila*. *Curr Biol.* 25(4):455–466.

Dyson CJ, Goodisman MAD. 2020. Gene duplication in the honeybee: patterns of DNA methylation, gene expression, and genomic environment. *Mol Biol Evol.* 37(8):2322–2331.

Ebrahim SAM, Dweck HKM, Stökl J, Hofferberth JE, Trona F, Weniger K, Rybak J, Seki Y, Stensmyr MC, Sachse S, et al. 2015. *Drosophila* avoids parasitoids by sensing their semiochemicals via a dedicated olfactory circuit. *PLoSBiol.* 13(12):e1002318.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32(5):1792–1797.

Eirín-López JM, Rebordinos L, Rooney AP, Rozas J. 2012. The birth-and-death evolution of multigene families revisited. *Genome Dyn*. 7:170–196.

Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol*. 16(1):157.

Engsontia P, Sangket U, Robertson HM, Satasook C. 2015. Diversification of the ant odorant receptor gene family and positive selection on candidate cuticular hydrocarbon receptors. *BMC Res Notes*. 8(1):380.

Espelie KE, Gamboa GJ, Grudzien BA, Bura EA. 1994. Cuticular hydrocarbons of the paper wasp, *Polistes fuscatus*: a search for recognition pheromones. *J Chem Ecol*. 20(7):1677–1687.

Espelie KE, Wenzel JW, Chang G. 1990. Surface lipids of social wasp *Polistes metricus* Say and its nest pedicel and their relation to nestmate recognition. *J Chem Ecol*. 16(7):2229–2241.

Ferguson ST, Park KY, Ruff AA, Bakis I, Zwiebel LJ. 2020. Odor coding of nestmate recognition in the eusocial ant *Camponotus floridanus*. *J Exp Biol*. 223(2):jeb215400.

Fishilevich E, Domingos AI, Asahina K, Naef F, Vosshall LB, Louis M. 2005. Chemotaxis behavior mediated by single larval olfactory neurons in *Drosophila*. *Curr Biol*. 15(23):2086–2096.

Fishilevich E, Vosshall LB. 2005. Genetic and functional subdivision of the *Drosophila* antennal lobe. *Curr Biol*. 15(17):1548–1553.

Fleischer J, Pregitzer P, Breer H, Krieger J. 2018. Access to the odor world: olfactory receptors and their role for signal transduction in insects. *Cell Mol Life Sci*. 75(3):485–508.

Fox AN, Pitts RJ, Robertson HM, Carlson JR, Zwiebel LJ. 2001. Candidate odorant receptors from the malaria vector mosquito *Anopheles gambiae* and evidence of down-regulation in response to blood feeding. *Proc Natl Acad Sci USA*. 98(25):14693–14697.

Freeman AR, Ophir AG, Sheehan MJ. 2020. The giant pouched rat (*Cricetomys ansorgei*) olfactory receptor repertoire. *PLoSOne*. 15(4):e0221981.

Galizia CG, Rössler W. 2010. Parallel Olfactory Systems in Insects: Anatomy and Function. *Ann Rev Entomol*. 55(1):399–420.

Gamboa GJ, Grudzien TA, Espelie KE, Bura EA. 1996. Kin recognition in social wasps: combining chemical and behavioural evidence. *AnimBehav*. 51(3):625–629.

Gamboa GJ, Reeve HK, Pfennig DW. 1986. The evolution and ontogeny of nestmate recognition in social wasps. *Annu Rev Entomol*. 31(1):431–454.

Goldman-Huertas B, Mitchell RF, Lapoint RT, Faucher CP, Hildebrand JG, Whiteman NK. 2015. Evolution of herbivory in Drosophilidae linked to loss of behaviors, antennal responses, odorant receptors, and ancestral diet. *Proc Natl Acad Sci Usa*. 112(10):3026–3031.

Greene MJ, Gordon DM. 2003. Cuticular hydrocarbons inform task decisions. *Nature*. 423(6935):32–32.

Groothuis J, Pfeiffer K, el Jundi B, Smid HM. 2019. The jewel wasp standard brain: average shape atlas and morphology of the female *Nasonia vitripennis* brain. *Arthropod Struct Dev*. 51:41–51.

Guo S, Kim J. 2007. Molecular evolution of *Drosophila* odorant receptor genes. *Mol Biol Evol*. 24(5):1198–1207.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc*. 8(8):1494–1512.

Hallem EA, Carlson JR. 2006. Coding of odors by a receptor repertoire. *Cell*. 125(1):143–160.

Hallem EA, Ho MG, Carlson JR. 2004. The molecular basis of odor coding in the *Drosophila* Antenna. *Cell*. 117(7):965–979.

Hansson BS, Stensmyr MC. 2011. Evolution of insect olfaction. *Neuron*. 72(5):698–711.

Harrop TWR, Guhlin J, McLaughlin GM, Permina E, Stockwell P, Gilligan J, Le Lec MF, Gruber MAM, Quinn O, Lovegrove M. 2020. High-quality assemblies for three invasive social wasps from the *Vespula* genus. *G3-Genes Genom Genet*. 10:3479–3488.

Hayden S, Bekaert M, Crider TA, Mariani S, Murphy WJ, Teeling EC. 2010. Ecological adaptation determines functional mammalian olfactory subgenomes. *Genome Res*. 20(1):1–9.

Hill CA, Fox NA, Pitts RJ, Kent LB, Tan PL, Chrystal MA, Cravchik A, Collins FH, Robertson HM, Zwiebel LJ. 2002. G protein-coupled receptors in *Anopheles gambiae*. *Science*. 298(5591):176–178.

Hines HM, Hunt JH, O'Connor TK, Gillespie JJ, Cameron SA. 2007. Multigene phylogeny reveals eusociality evolved twice in vespid wasps. *Proc Natl Acad Sci USA*. 104(9):3295–3299.

Holt C, Yandell M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics*. 12:491.

Hopf TA, Morinaga S, Ihara S, Touhara K, Marks DS, Benton R. 2015. Amino acid coevolution reveals three-dimensional structure and functional domains of insect odorant receptors. *Nat Commun*. 6(1):6077.

Jandt JM, Tibbetts EA, Toth AL. 2014. *Polistes* paper wasps: a model genus for the study of social dominance hierarchies. *Insect Soc*. 61(1):11–27.

Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*. 8(3):275–282.

Käll L, Krogh A, Sonnhammer ELL. 2004. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol*. 338(5):1027–1036.

Kapheim KM, Pan H, Li C, Salzberg SL, Puiu D, Magoc T, Robertson HM, Hudson ME, Venkat A, Fischman BJ, et al. 2015. Genomic signatures of evolutionary transitions from solitary to group living. *Science*. 348(6239):1139–1143.

Karpe SD, Dhingra S, Brockmann A, Sowdhamini R. 2017. Computational genome-wide survey of odorant receptors from two solitary bees *Dufourea novaeangliae* (Hymenoptera: halictidae) and *Habropoda laboriosa*. *Sci Rep*. 7(1):10823.

Karpe SD, Jain R, Brockmann A, Sowdhamini R. 2016. Identification of complete repertoire of *Apisflorea*odorant receptors reveals complex orthologous relationships with *Apis mellifera*. *Genome Biol Evol*. 8(9):2879–2895.

Karpe SD, Tiwari V, Ramanathan S. 2021. InsectOR—Webserver for sensitive identification of insect olfactory receptor genes from non-model genomes. *PLoS One* 16(1):e0245324.

Kather R, Martin SJ. 2015. Evolution of cuticular hydrocarbons in the Hymenoptera: a meta-analysis. *J Chem Ecol*. 41(10):871–883.

Keller A, Vosshall LB. 2007. Influence of odorant receptor repertoire on odor perception in humans and fruit flies. *Proc Natl Acad Sci USA*. 104(13):5614–5619.

Kucharski R, Maleszka J, Maleszka R. 2016. A possible role of DNA methylation in functional divergence of a fast evolving duplicate gene encoding odorant binding protein 11 in the honeybee. *Proc R Soc B*. 283(1833):20160558.

Lavine BK, Morel L, Vander Meer RK, Gunderson RW, Han JH, Bonanno A, Stine A. 1990. Pattern recognition studies in chemical communication: nestmate recognition in *Camponotus floridanus*. *ChemometrIntell Lab*. 9(1):107–114.

Layton JM, Camann MA, Espelie KE. 1994. Cuticular lipid profiles of queens, workers, and males of social wasp *Polistes metricus* Say are colony-specific. *J Chem Ecol*. 20(9):2307–2321.

LeBoeuf AC, Benton R, Keller L. 2013. The molecular basis of social behavior: models, methods and advances. *Curr Opin Neurobiol*. 23(1):3–10.

Leonhardt SD, Menzel F, Nehring V, Schmitt T. 2016. Ecology and evolution of communication in social insects. *Cell*. 164(6):1277–1287.

Lewis WJ, Sonnet PE, Nordlund DA. 1988. Responses of braconid parasitoids *Microplitis croceipes* (Cresson) and *M. demolitor* Wilkonson to stereoisomers of kairomone 13-methylhentriacontane. *J Chem Ecol*. 14(3):883–888.

Lynch M. 2007. The origins of genome architecture. Sunderland (MA): Sinauer Associates.

Malnic B, Hirono J, Sato T, Buck LB. 1999. Combinatorial receptor codes for odors. *Cell*. 96(5):713–723.

Mansourian S, Stensmyr MC. 2015. The chemical ecology of the fly. *Curr Opin Neurobiol*. 34:95–102.

Masson C, Strambi C. 1977. Sensory antennal organization in an ant and a wasp. *J Neurobiol*. 8(6):537–548.
Mathew D, Martelli C, Kelley-Swift E, Brusalis C, Gershow M, Samuel ADT, Emonet T, Carlson JR. 2013. Functional diversity among sensory receptors in a *Drosophila* olfactory circuit. *Proc Natl Acad Sci USA*. 110(23):E2134–E2143.
McBride CS. 2007. Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*. *Proc Natl Acad Sci USA*. 104(12):4996–5001.
McBride CS, Arguello JR. 2007. Five *Drosophila* genomes reveal nonneutral evolution and the signature of host specialization in the chemoreceptor superfamily. *Genetics*. 177(3):1395–1416.
McBride CS, Baier F, Omondi AB, Spitzer SA, Lutomiah J, Sang R, Ignell R, Vosshall LB. 2014. Evolution of mosquito preference for humans linked to an odorant receptor. *Nature*. 515(7526):222–227.
McKenzie SK, Fetter-Pruneda I, Ruta V, Kronauer DJC. 2016. Transcriptomics and neuroanatomy of the clonal raider ant implicate an expanded clade of odorant receptors in chemical communication. *Proc Natl Acad Sci USA*. 113(49):14091–14096.
McKenzie SK, Kronauer DJC. 2018. The genomic architecture and molecular evolution of ant odorant receptors. *Genome Res*. 28(11):1757–1765.
Miller CH, Campbell P, Sheehan MJ. 2020. Distinct evolutionary trajectories of V1R clades across mouse species. *BMC Evol Biol*. 20(1):99.
Miller SE, Legan AW, Flores ZA, Ng HY, Sheehan MJ. 2019. Strong, but incomplete, mate choice discrimination between two closely related species of paper wasp. *Biol J Linn Soc*. 126(3):614–622.
Miller SE, Legan AW, Henshaw M, Ostevik KL, Samuk K, Uy FMK, Sheehan MJ. 2020. Evolutionary dynamics of recent selection on cognitive abilities. *Proc Natl Acad Sci Usa*. 117(6):3045–3052.
Miller SE, Bluher SE, Bell E, Cini A, Silva R. C D,de Souza AR, Gandia KM, Jandt J, Loope K, Prato A, et al. 2018. WASPnest: a worldwide assessment of social Polistine nesting behavior. *Ecology*. 99(10):2405–2405,,.
Münch D, Galizia CG. 2016. DoOR 2.0 - comprehensive mapping of *Drosophila melanogaster* odorant responses. *Sci Rep*. 6(1):21841.
Nagawa F, Yoshihara S-I, Tsuboi A, Serizawa S, Itoh K, Sakano H. 2002. Genomic analysis of the murine odorant receptor *MOR28* cluster: a possible role of gene conversion in maintaining the olfactory map. *Gene*. 292(1-2):73–80.
Nei M. 2007. The new mutation theory of phenotypic evolution. *Proc Natl Acad Sci USA*. 104(30):12235–12242.
Nei M. 2013. Mutation-driven evolution. Oxford (UK): Oxford University Press.
Niehuis O, Büllesbach J, Judson AK, Schmitt T, Gadau J. 2011. Genetics of cuticular hydrocarbon differences between males of the parasitoid wasps *Nasonia giraulti* and *Nasonia vitripennis*. *Heredity*. 107(1):61–70.
Nozawa M, Nei M. 2007. Evolutionary dynamics of olfactory receptor genes in *Drosophila* species. *Proc Natl Acad Sci USA*. 104(17):7122–7127.
O'Neill KM. 2001. Solitary wasps: behavior and natural history. Ithaca (NY): Cornell University Press.
Ohno S. 1970. Evolution by gene duplication. Berlin: Springer.
Oi CA, Oliveira RC, van Zweden JS, Mateus S, Millar JG, Nascimento FS, Wenseleers T. 2019. Do primitively eusocial wasps use queen pheromones to regulate reproduction? A case study of the paper wasp *Polistes satan*. *Front EcolEvol*. 7:199.
Page RE, Metcalf RA, Metcalf RL, Erickson EH, Lampman RL. 1991. Extractable hydrocarbons and kin recognition in honeybee (*Apis mellifera* L.). *J Chem Ecol*. 17(4):745–756.
Pask GM, Slone JD, Millar JG, Das P, Moreira JA, Zhou X, Bello J, Berger SL, Bonasio R, Desplan C, et al. 2017. Specialized odorant receptors in social insects that detect cuticular hydrocarbon cues and candidate pheromones. *Nat Commun*. 8(1):297.
Patalano S, Vlasova A, Wyatt C, Ewels P, Camara F, Ferreira PG, Asher CL, Jurkowski TP, Segonds-Pichon A, Bachman M, et al. 2015. Molecular signatures of plastic phenotypes in two eusocial insect species with simple societies. *Proc Natl Acad Sci USA*. 112(45):13970–13975.
Peeters C, Liebig J. 2009. Organization of insect societies: from genome to socio-complexity. Fertility signaling as a general mechanism of regulating reproductive division of labor in ants. In: Gadau J, Fewell J, Wilson EO, editors. Cambridge (MA): Harvard University Press. p. 220–242.
Peters RS, Krogmann L, Mayer C, Donath A, Gunkel S, Meusemann K, Kozlov A, Podsiadlowski L, Petersen M, Lanfear R, et al. 2017. Evolutionary History of the Hymenoptera. *Curr Biol*. 27(7):1013–1018.
Pond SLK, Poon AFY, Velazquez R, Weaver S, Hepler NL, Murrell B, Shank SD, Magalis BR, Bouvier D, Nekrutenko A, et al. 2020. HyPhy 2.5 – A customizable platform for evolutionary hypothesis testing using phylogenies. *Mol Biol Evol*. 37(1):295–299.
Post DC, Jeanne RL. 1984. Recognition of conspecifics and sex by territorial males of the social wasp *Polistes fuscatus* (Hymenoptera: Vespidae). *Behav*. 91(1-3):78–92.
Post DC, Jeanne RL. 2010. Male reproductive behavior of the social wasp *Polistes fuscatus* (Hymenoptera: Vespidae). *Z Tierpsychol*. 62(2):157–171.
Rabb RL. 1960. Biological studies of *Polistes* in North Carolina (Hymenoptera: Vespidae). *Ann Entomol Soc Am*. 53(1):111–121.
Rabb RL, Lawson FR. 1957. Some factors influencing the predation of *Polistes* wasps on the tobacco hornworm. *J Econ Entomol*. 50(6):778–784.
Raguso RA. 2008. Wake up and smell the roses: the ecology and evolution of floral scent. *Annu Rev Ecol Evol Syst*. 39(1):549–569.
Ramdya P, Benton R. 2010. Evolving olfactory systems on the fly. *Trends Genet*. 26(7):307–316.
Reed HC, Landolt PJ. 1990. Sex attraction in paper wasp, *Polistes exclamans* Viereck (Hymenoptera: vespidae), in a wind tunnel. *J Chem Ecol*. 6(4):1277–1287.
Reeve HK. 1991. Polistes. In: Ross KG, Matthews RD, editors. The social biology of wasps. Ithaca (NY): Cornell University Press. p. 99–148.
Rehan SM, Toth AL. 2015. Climbing the social ladder: the molecular evolution of sociality. *Trends Ecol Evol*. 30(7):426–433.
Richter MR. 2000. Social wasp (Hymenoptera: Vespidae) foraging behavior. *Annu Rev Entomol*. 45(1):121–150.
Robertson HM, Gadau J, Wanner KW. 2010. The insect chemoreceptor superfamily of the parasitoid jewel wasp *Nasonia vitripennis*. *Insect Mol Biol*. 19:121–136.
Robertson HM, Wanner KW. 2006. The chemoreceptor superfamily in the honeybee *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family. *Genome Res*. 16(11):1395–1403.
Robertson HM, Warr CG, Carlson JR. 2003. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc Natl Acad Sci USA*. 100(Supplement 2):14537–14542.
Ross KG, Matthews RW. 1991. The social biology of wasps. Ithaca (NY): Cornell University Press.
Roux J, Privman E, Moretti S, Daub JT, Robinson-Rechavi M, Keller L. 2014. Patterns of positive selection in seven ant genomes. *Mol Biol Evol*. 31(7):1661–1685.
Saad R, Cohanim AB, Kosloff M, Privman E. 2018. Neofunctionalization in ligand binding sites of ant olfactory receptors. *Genome Biol Evol*. 10(9):2490–2500.
Sadd BM, Barribeau SM, Bloch G, de Graaf DC, Dearden P, Elsik CG, Gadau J, Grimmelikhuijzen CJP, Hasselmann M, Lozier JD, et al. 2015. The genomes of two key bumblebee species with primitive eusocial organization. *Genome Biol*. 16(1):76.
Sánchez-Gracia A, Vieira FG, Rozas J. 2009. Molecular evolution of the major chemosensory gene families in insects. *Heredity*. 103(3):208–216.
Sato KM, Pellegrino M, Nakagawa T, Nakagawa T, Vosshall LB, Touhara K. 2008. Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature*. 452(7190):1002–1006.
Sawyer SA. 1989. Statistical tests for detecting gene conversion. *Mol Biol Evol*. 6(5):526–538.

Seeholzer LF, Seppo M, Stern DL, Ruta V. 2018. Evolution of a central neural circuit underlies *Drosophila* mate preferences. *Nature*. 559(7715):564–569.

Sheehan MJ, Botero CA, Hendry TA, Sedio BE, Jandt JM, Weiner S, Toth AL, Tibbetts EA. 2015. Different axes of environmental variation explain the presence vs. extent of cooperative nest founding associations in *Polistes* paper wasps. *Ecol Lett*. 18(10):1057–1067.

Singer TL. 1998. Roles of hydrocarbons in the recognition systems of insects. *Am Zool*. 38(2):394–405.

Singer TL, Camann MA, Espelie KE. 1992. Discriminant analysis of cuticular hydrocarbons of social wasp *Polistes exclamans* Viereck and surface hydrocarbons of its nest paper and pedicel. *J Chem Ecol*. 18(5):785–797.

Sledge MF, Boscaro F, Turillazzi S. 2001a. Cuticular hydrocarbons and reproductive status in the social wasp *Polistes dominulus*. *Behav Ecol Sociobiol*. 49(5):401–409.

Sledge MF, Dani FR, Cervo R, Dapporto L, Turillazzi S. 2001b. Recognition of social parasites as nest-mates: adoption of colony-specific host cuticular odours by the paper wasp parasite *Polistes sulcifer*. *Proc R Soc Lond B*. 268(1482):2253–2260.

Sledge MF, Trinca I, Massolo A, Boscaro F, Turillazzi S. 2004. Variation in cuticular hydrocarbon signatures, hormonal correlates and establishment of reproductive dominance in a polistine wasp. *J Insect Physiol*. 50(1):73–83.

Slone JD, Pask GM, Ferguson ST, Millar JG, Berger SL, Reinberg D, Liebig J, Ray A, Zwiebel LJ. 2017. Functional characterization of odorant receptors in the ponerine ant, *Harpegnathos saltator*. *Proc Natl Acad Sci Usa*. 114(32):8586–8591.

Smith CR, Smith CD, Robertson HM, Helmkampf M, Zimin A, Yandell M, Holt C, Hu H, Abouheif E, Benton R, et al. 2011. Draft genome of the red harvester ant *Pogonomyrmex barbatus*. *Proc Natl Acad Sci USA*. 108(14):5667–5672.

Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Pond SLK. 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Genome Biol Evol*. 32(5):1342–1353.

Smith CD, Zimin A, Holt C, Abouheif E, Benton R, Cash E, Croset V, Currie CR, Elhaik E, Elsik CG, et al. 2011. Draft genome of the globally widespread and invasive Argentine ant (*Linepithema humile*). *Proc Natl Acad Sci USA*. 108(14):5673–5678.

Sonnhammer ELL, von Heijne G, Krogh A. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol*. 6:175–182.

Southon RJ, Fernandes OA, Nascimento FS, Sumner S. 2019. Social wasps are effective biocontrol agents of key lepidopteran crop pests. *Proc R Soc B*. 286(1914):20191676.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9):1312–1313.

Standage DS, Berens AJ, Glastad KM, Severin AJ, Brendel VP, Toth AL. 2016. Genome, transcriptome and methylome sequencing of a primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social insect. *Mol Ecol*. 25(8):1769–1784.

Stensmyr MC, Dweck HKM, Farhan A, Ibba I, Strutz A, Mukunda L, Linz J, Grabe V, Steck K, Lavista-Llanos S, et al. 2012. A conserved dedicated olfactory circuit for detecting harmful microbes in *Drosophila*. *Cell*. 151(6):1345–1357.

Stuart RJ. 1988. Collective cues as a basis for nestmate recognition in polygynous leptothoracine ants. *Proc Natl Acad USA*. 85(12):4572–4575.

Sturgis SJ, Gordon DM. 2012. Nestmate recognition in ants (Hymenoptera: formicidae): a review. *Myrmecol News*. 16:101–110.

Su C-Y, Menuz K, Carlson JR. 2009. Olfactory perception: receptors, cells, and circuits. *Cell*. 139(1):45–59.

Suh E, Bohbot JD, Zwiebel LJ. 2014. Peripheral olfactory signaling in insects. *Curr Opin Insect Sci*. 6:86–92.

Toth AL, Tooker JF, Radhakrishnan S, Minard R, Henshaw MT, Grozinger CM. 2014. Shared genes related to aggression, rather than chemical communication, are associated with reproductive dominance in paper wasps (*Polistes metricus*). *BMC Genomics*. 15:75.

Touhara K, Vosshall L. 2009. Sensing odorants and pheromones with chemosensory receptors. *Annu Rev Physiol*. 71(1):307–332.

Trible W, Olivos-Cisneros L, McKenzie SK, Saragosti J, Chang N-C, Matthews BJ, Oxley PR, Kronauer DJ. 2017. Orco mutagenesis causes loss of antennal lobe glomeruli and impaired social behavior in ants. *Cell*. 170(4):727–735.

Tsutsui ND. 2013. Dissecting ant recognition systems in the age of genomics. *Biol Lett*. 9(6):20130416.

van Zweden JS, d'Ettorre P. 2010. Nestmate recognition in social insects and the role of hydrocarbons. In: Blomquist GJ, Bagnères A-G, editors. Insect Hydrocarbons. Cambridge (UK): Cambridge University Press. p. 222–243.

Vernot B, Stolzer M, Goldman A, Durand D. 2008. Reconciliation with Non-Binary Species Trees. *J Comput Biol*. 15(8):981–1006.

Vizueta J, Escuer P, Frías-López C, Guirao-Rico S, Hering L, Mayer G, Rozas J, Sánchez-Gracia A. 2020. Evolutionary history of major chemosensory gene families across Panarthropoda. *Mol Biol Evol*. msaa197.

Vogt RG, Prestwich GD, Lerner MR. 1991. Odorant-binding-protein subfamilies associate with distinct classes of olfactory receptor neurons in insects. *J Neurobiol*. 22(1):74–84.

Vosshall LB, Amrein H, Morozov PS, Rzhetsky A, Axel R. 1999. A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell*. 96(5):725–736.

Wanner KW, Nichols AS, Walden KKO, Brockmann A, Luetje CW, Robertson HM. 2007. A honey bee odorant receptor for the queen substance 9-oxo-2-decenoic acid. *Proc Natl Acad Sci USA*. 104(36):14383–14388.

Wertheim JO, Murrell B, Smith MD, Pond SLK, Scheffler K. 2015. RELAX: detecting relaxed selection in a phylogenetic framework. *Mol Biol Evol*. 32(3):820–832.

Wicher D. 2015. Molecular basis of olfaction. In: Glatz R, editor. Olfactory signaling in insects. Elsevier. p. 37–54.

Xiao J-H, Yue Z, Jia L-Y, Yang X-H, Niu L-H, Wang Z, Zhang P, Sun B-F, He S-M, Li Z, et al. 2013. Obligate mutualism within a host drives the extreme specialization of a fig wasp genome. *Genome Biol*. 14(12):R141.

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

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 24(8):1586–1591.

Yang Z, Nielsen R. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol*. 17(1):32–43.

Yan H, Opachaloemphan C, Mancini G, Yang H, Gallitto M, Mlejnek J, Leibholz A, Haight K, Ghaninia M, Huo L, et al. 2017. An engineered orco mutation produces aberrant social behavior and defective neural development in ants. *Cell*. 170(4):736–747.,

Yapici N, Zimmer M, Domingos AI. 2014. Cellular and molecular basis of decision-making. *EMBO Rep*. 15(10):1023–1035.

Young JM, Friedman C, Williams EM, Ross JA, Tonnes-Priddy L, Trask BJ. 2002. Different evolutionary processes shaped the mouse and human olfactory receptor gene families. *Hum Mol Genet*. 11(5):535–546.

Zhou X, Rokas A, Berger SL, Liebig J, Ray A, Zwiebel LJ. 2015. Chemoreceptor evolution in hymenoptera and its implications for the evolution of eusociality. *Genome Biol Evol*. 7(8):2407–2416.

Zhou X, Slone JD, Rokas A, Berger SL, Liebig J, Ray A, Reinberg D, Zwiebel LJ. 2012. Phylogenetic and transcriptomic analysis of chemosensory receptors in a pair of divergent ant species reveals sex-specific signatures of odor coding. *PLoSGenet*. 8(8):e1002930.

Zube C, Rössler W. 2008. Caste- and sex-specific adaptations within the olfactory pathway in the brain of the ant *Camponotus floridanus*. *Arthropod Struct Dev*. 37(6):469–479.