Expansion and accelerated evolution of 9-exon odorant receptors in *Polistes* paper wasps

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

a. Department of Neurobiology and Behavior, Cornell University, Ithaca, NY, USA
*Authors for Correspondence: Andrew W. Legan: awl75@cornell.edu; Michael J. Sheehan: msheehan@cornell.edu

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 odorant receptor gene family evolution in the northern paper wasp, *Polistes fuscatus*, and compare it to four additional paper wasp species spanning ~40 million years of divergence. We find 200 functional OR genes in *P. fuscatus* matching predictions from neuroanatomy, and more than half of these are in the 9-exon subfamily. Lineage-specific expansions of 9-exon subfamily ORs are tandemly arrayed in *Polistes* genomes and exhibit a breakdown in microsynteny relative to tandem arrays in other OR subfamilies. There is evidence of episodic positive diversifying selection shaping ORs in expanded subfamilies, including 9-exon, E, H, and L, but 9-exon ORs do not stand out as selectively diversified among *Polistes* species. Accelerated evolution has resulted in lower amino acid similarity and high d_N/d_S among 9-exon ORs compared to other OR subfamilies. Patterns of OR evolution within *Polistes* are consistent with 9-exon OR function in CHC perception by combinatorial coding, with both selection and drift contributing to interspecies differences in copy number and sequence.

Keywords: birth-and-death process, comparative genomics, tandem array, olfaction, social insect, antennal lobe glomeruli

INTRODUCTION

Odorant receptors are among the largest gene families in many animal genomes, with variation in the relative number of genes thought to reflect aspects of species chemosensory ecology. From the standpoint of molecular evolution, the odorant receptors (ORs) of insects and olfactory receptors in vertebrates have been widely studied as a model to understand the dynamics of gene family evolution (Nozawa & Nei 2007; Nei 2013; Benton 2015). Yet fundamental features of odorant receptor 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 OR gene family at broad phylogenetic scales, where there is often little 1:1 orthology of receptors among species (Tsutsui 2013; Yan et al. 2020). At the same time, studies of closely related species with similar ecologies and life histories can reveal the dynamics of receptor evolution at finer timescales and elucidate the process of gene family turnover (Guo & Kim 2007; Brand et al. 2015; Karpe et al. 2016; Brand & Ramirez 2017; 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; Engsontia et al. 2015; Kapheim et al. 2015; Karpe et al. 2016, 2017; McKenzie et al. 2016). 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 odorant receptor 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, studies of social insect ORs have focused on comparisons between genera or families, meaning the dynamics of OR gene family turnover across a gradient of species divergence times remain to be investigated. A better understanding of the short-term mechanisms of OR evolution provides additional insights

57 into the molecular evolutionary dynamics shaping receptor diversity across more distantly related
58 taxa. The recent release of five genomes of *Polistes* paper wasps spanning ~1-40 million years of
59 divergence provide an opportunity to fill these gaps in our knowledge.

60 The molecular evolution of the OR gene family is best described as a birth-and-death process, in
61 which genes are duplicated and deleted over evolutionary time (Nei 2007; Nozawa & Nei 2007;
62 Eirín-López et al. 2012). Both random drift and natural selection are present during this process,
63 determining the extent of OR gene copy number variation and the rate of gene sequence evolution
64 (Nei 2007; Nozawa & Nei 2007). It is useful to subdivide odorant receptor function into behavioral
65 and molecular functions, in order to examine how OR function at the organismal and molecular
66 levels corresponds to different patterns of molecular evolution. Organisms use odorant receptor
67 proteins to detect stimuli and provide input to neural circuits that regulate decision-making, here
68 termed behavioral function (Yapici et al. 2014). At the molecular level, the specificity with which
69 an OR binds chemical compounds or families of compounds (ligands) determines the tuning
70 specificity of the olfactory receptor neuron (ORN) in which it is expressed (Hallem et al. 2004).
71 ORN tuning specificity describes the relationship between the change in the frequency of action
72 potentials fired by the ORN across a range of different chemical stimuli. An OR's molecular
73 function can be categorized as specialist if ligand binding is highly specific, i.e. one chemical, or
74 generalist if ligand binding allows for detection of multiple chemicals or chemical types. Narrowly
75 tuned ORNs express specialist ORs that bind a specific ligand, creating a dedicated channel of
76 olfaction, while broadly tuned ORNs express generalist ORs that bind a larger spectrum of ligands
77 (Touhara & Vosshall 2009; Andersson et al. 2015). In some cases, ORNs achieve broad tuning by
78 co-expressing multiple types of ORs (Fleischer et al. 2018). Generalist ORs are important for
79 combinatorial coding: the process of combining input from multiple ORs that bind overlapping
80 sets of ligands in order to discriminate a large variety of odors (Malnic et al. 1999). In nature, ORs
81 exist on a spectrum from specialist to generalist, and ORN responses are dependent upon a
82 complex interaction of OR, ligand concentration, and odorant binding proteins (Vogt et al. 1991;
83 Hallem et al. 2004; Hallem & Carlson 2006; Stensmyr et al. 2012; Mathew et al. 2013; Dweck et
84 al. 2015; Ebrahim et al. 2015; Münch & Galizia 2016).

85 There are distinct sets of predictions for the molecular evolution of ORs depending on their
86 behavioral and molecular functions. Negative selection (purifying selection) is expected to
87 conserve specialist ORs, which are not resilient to mutations due to the tight link between receptor
88 and ligand in a dedicated olfactory channel (Andersson et al. 2015). For example, a volatile emitted
89 by toxic microbes triggers avoidance behavior in *Drosophila melanogaster* by binding a specialist
90 OR that is conserved across the genus (Stensmyr et al. 2012). In contrast to dedicated olfactory
91 channels, combinatorial coding is thought to involve ORs with overlapping responses to ligands
92 (Andersson et al. 2015). Mutations that slightly alter the response profiles of functionally
93 redundant ORs may not be eliminated by negative selection, since other ORs can help compensate
94 (Fishilevich et al. 2005; Keller & Vosshall 2007). Copy number variation and relaxed selection
95 allow ORs to gain mutations that might endow them with an adaptive behavioral function.
96 Divergent chemosensory landscapes between species lead to increased copy number variation as
97 ORs with new behavioral functions are gained and ORs that detect irrelevant ligands are lost
98 (Ramdya & Benton 2010; Goldman-Huertas et al. 2015).

99 The molecular evolution of the OR gene family is dynamic among the Hymenoptera, with
100 prevalent lineage-specific gene expansions and losses, especially in the 9-exon OR subfamily
101 (Engsontia et al. 2015; Zhou et al. 2015; McKenzie & Kronauer 2018). The 9-exon ORs constitute
102 about one third of all ant ORs, and have evolved rapidly in ants, frequently under positive selection,
103 leading researchers to propose that 9-exon ORs facilitate recognition of cuticular hydrocarbons
104 (CHCs) (Smith CR, Smith CD et al. 2011; Smith CD, Zimin et al. 2011; Zhou et al. 2012, 2015;
105 Engsontia et al. 2015; McKenzie et al. 2016). CHCs are used by insects to waterproof the cuticle
106 and to communicate with conspecifics (Blomquist & Bagnères 2010). While less pronounced than
107 in ants, dynamic evolution is also characteristic of 9-exon OR evolution in social bees, which rely
108 on CHCs in communication (Sadd et al. 2015; Karpe et al. 2016, 2017). Functional studies in
109 which ORs were transfected into an empty *D. melanogaster* ORN have verified that at least some
110 9-exon ORs of the ant *Harpegnathos saltator* overlap in their responses to ligands, with multiple
111 9-exon ORs responding to the same CHC molecule and unique 9-exon ORs responding to multiple
112 different CHC molecules (Pask et al. 2017; Slone et al. 2017). Functional ORs are necessary for
113 normal nesting behavior and for nestmate recognition in ants, a process which involves detecting
114 variation in the CHCs on the cuticles of conspecifics (Lavine et al. 1990; van Zweden & d'Ettorre

115 2010; Sturgis & Gordon 2012; Trible et al. 2017; Yan et al. 2017; Ferguson et al. 2020). Together
116 these studies suggest that 9-exon ORs function in combinatorial coding of CHC perception.

117 Like other Hymenopterans, vespid wasps, including the genus *Polistes*, use CHCs in complex
118 social behaviors (Gamboa et al. 1986, 1996; Dani & Turillazzi 2018). *Polistes* use chemicals as
119 signals and cues in a variety of behaviors, including during mate attraction, mate compatibility
120 recognition, queen recognition, dominance/fertility signaling, and nestmate recognition (Reed &
121 Landolt 1990; Espelie et al. 1994; Post & Jeanne 1984; Dapporto et al. 2007; Jandt et al. 2014;
122 Sledge et al. 2001a, b, 2004; Oi et al. 2019). Recent efforts to sequence *Polistes* genomes provide
123 an opportunity to resolve patterns of OR evolution among closely related species as an independent
124 test of 9-exon OR gene subfamily expansion during social evolution (Patalano et al. 2015;
125 Standage et al. 2016; Miller SE et al. 2020). We annotated the OR repertoires of five *Polistes*
126 species representing ~40 million years of evolution: *P. fuscatus*, *P. metricus*, *P. dorsalis*, *P.*
127 *canadensis*, and *P. dominula* (Figure 1). Combining neuroanatomy, manual gene annotation, and
128 molecular evolution analysis, we examined the evolution of odorant receptors in light of the
129 patterns predicted for different behavioral and molecular functions. We discover that social wasps,
130 like ants, have an expanded subfamily of 9-exon ORs. Between *Polistes* species, 9-exon ORs
131 exhibit dynamic evolution relative to ORs in other subfamilies, which are highly conserved.
132 Expansion and birth-and-death evolution of the 9-exon OR subfamily in social wasps is consistent
133 with a unique function in combinatorial coding perception of CHCs.

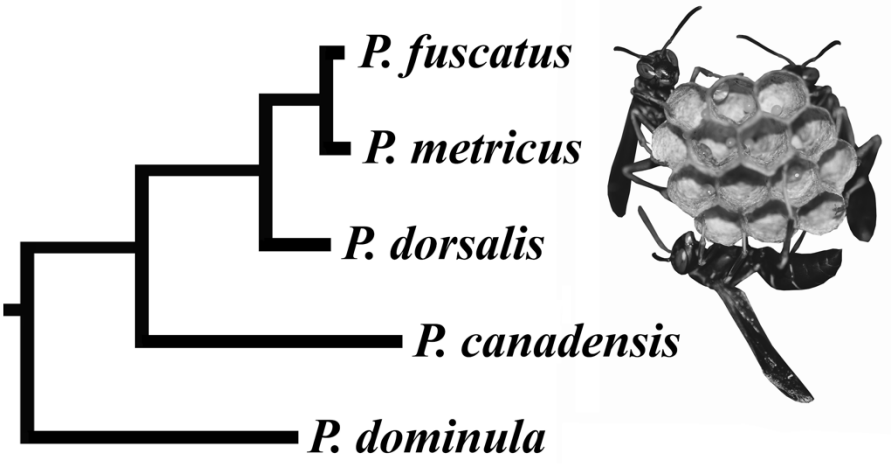

Fig. 1: Phylogeny of five *Polistes* species considered in this study: *P. fuscatus*, *P. metricus*, *P. dorsalis*, *P. canadensis,* and *P. dominula*. The photo to the right of the phylogeny shows *P. fuscatus* foundresses on a nest. Phylogenetic tree based on the 16S ribosomal RNA gene and the cytochrome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

147 **RESULTS**

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

149 In order to predict the OR repertoires of *P. fuscatus* and four other *Polistes* species, we combined
150 fluorescent confocal microscopy of the *P. fuscatus* antennal lobe with manual genome annotation
151 informed by antennal RNAseq. We found 229 glomeruli in the antennal lobe of an adult gyne
152 (female reproductive) (Figure S1). Across a sample of insects, the number of intact OR genes in
153 the genome correlates with the number of glomeruli in the antennal lobe, predicting 229 ORs in
154 the *P. fuscatus* genome (Figure 2). Here we focus on the *P. fuscatus* genome because it has nearly
155 chromosome level scaffolds and is the best assembled *Polistes* genome (Table S1; Patalano et al.
156 2015; Standage et al. 2016; Miller SE et al. 2020). Automated annotation using the MAKER
157 pipeline (Holt & Yandell 2011) without guidance from antennal mRNA predicted 115 OR gene
158 models in the *P. fuscatus* genome. A combined *P. fuscatus* male and gyne (reproductive female)
159 antennal transcriptome generated using Trinity (Haas et al. 2013) yielded 89 OR genes greater than
160 900 nucleotides in length. Some long Trinity genes contain multiple 7-transmembrane domains
161 and likely represent concatenated OR genes. The small fraction of the *P. fuscatus* OR repertoire
162 predicted by transcriptome assembly is consistent with previous observations that annotation of
163 OR repertoires using only transcriptome data typically fails to recover all ORs (Karpe et al. 2016,
164 2017, 2020).

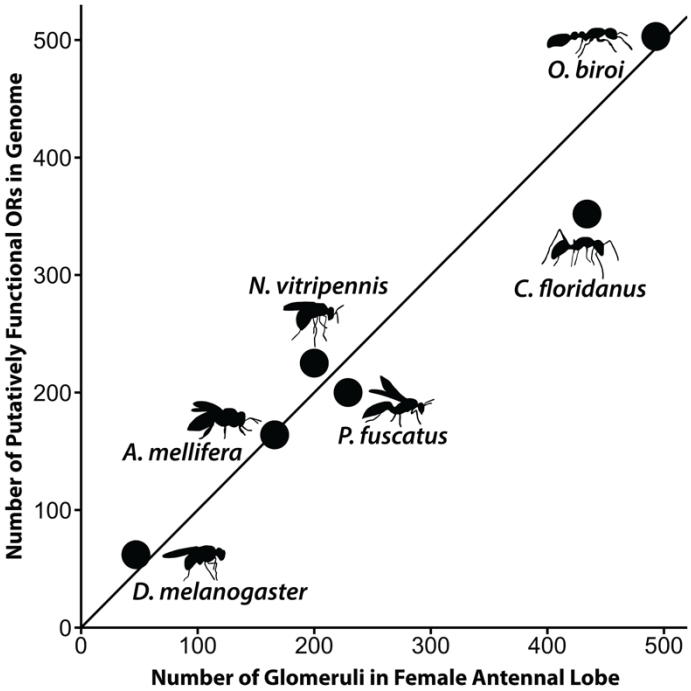

Fig. 2: The number of functional ORs is correlated with the number of antennal lobe glomeruli across insect species (50 glomeruli and 62 ORs in the genome of the common fruit fly *D. melanogaster*, Fishilevich & Vosshall 2005; 166 glomeruli in the worker and 163 intact ORs in the genome of the honey bee *A. mellifera*, Arnold et al. 1985, Robertson & 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 & 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 & Kronauer 2018). The diagonal line represents a line of equality with slope of 1.

185 Manual gene annotation of *P. fuscatus* ORs recovered 231 gene models across 28 scaffolds (Figure
186 S2), of which 28 are pseudogenes and 10 are incomplete gene models (7 missing N termini, 2
187 missing C termini, and one missing both N and C termini). Since functional insect ORs are
188 typically composed of 400 amino acids, we defined gene models as putatively functional if they
189 coded for proteins greater than or equal to 300 amino acids in length. In *P. fuscatus*, the 200
190 putatively functional gene models encode protein sequences with an average length of 395 ± 15
191 (SD) amino acids, and 198 of these gene models encode protein sequences greater than 350 amino
192 acids in length (Table 1). Odorant receptor proteins possess seven transmembrane domains
193 (Wicher 2015). The putatively functional *P. fuscatus* OR proteins possess on average 5.95 ± 0.91
194 (SD) transmembrane domains as predicted by TMHMM version 2.0c (Sonnhammer et al. 1998)
195 and 6.43 ± 1.13 (SD) as predicted by Phobius version 1.01 (Käll et al. 2004). For comparison,
196 transmembrane domain prediction in 61 *D. melanogaster* ORs coding for proteins greater than 375
197 amino acids in length found on average 5.77 ± 1.12 (SD) transmembrane domains as predicted by
198 TMHMM version 2.0c and 6.18 ± 1.09 (SD) as predicted by Phobius version 1.01 (sequences from
199 Hopf et al. 2015 Supplemental Data 1). The close match between the number of ORs predicted by
200 neuroanatomy and the number recovered from manual annotation suggests that we have identified
201 nearly all of the OR genes in *P. fuscatus*. The number of transmembrane domains predicted are
202 comparable to annotations of *D. melanogaster* and approach the 7 transmembrane domains
203 expected for insect ORs. Manual OR gene annotation in *P. fuscatus* and four other *Polistes*
204 genomes is summarized in Table 1.

205 **Table 1**: 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

206 a. Putatively functional ORs ≥ 300 amino acids in length; b. Mean length in amino acids (± SD); c. Mean
207 transmembrane domains predicted by TMHMM (± SD); d. Mean transmembrane domains predicted by
208 Phobius (± SD)

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 lineage-specific expansions of multiple OR subfamilies (Figure 3). Gene gain and loss events were predicted using NOTUNG (Chen et al. 2000) and mapped onto a species cladogram of 14 Hymenopterans (Figure 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 (Figure 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-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 (Figure 4). Given the well-documented use of CHCs as signal molecules in *Polistes* (Singer 1998; Dani et al. 2001; Dani 2009), it is not surprising to find expansions in the 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.

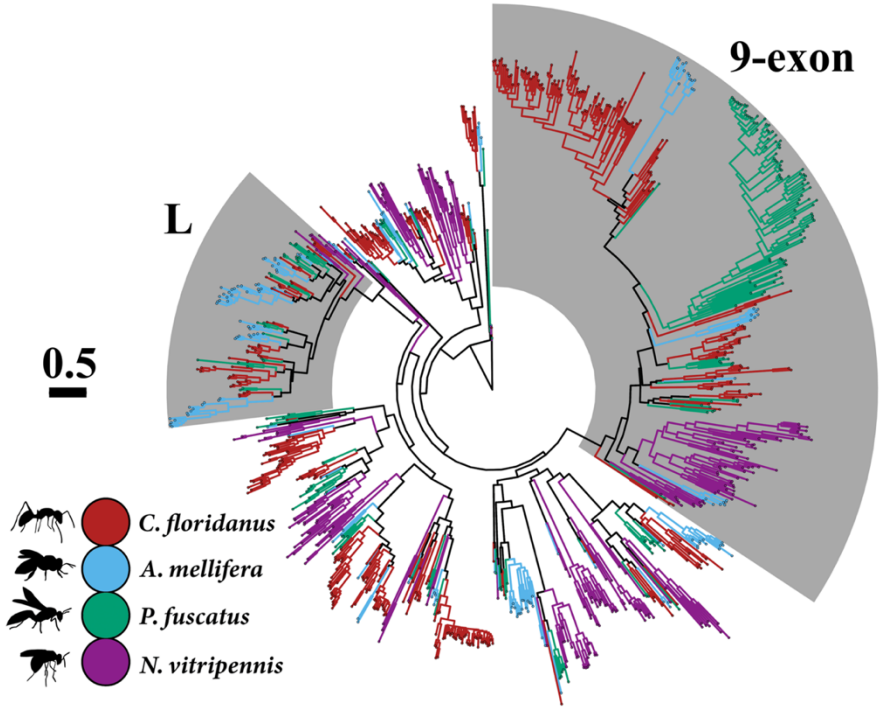

Fig. 3: Maximum likelihood OR protein tree constructed using data from four Hymenopterans (*Apis mellifera*, Robertson & Wanner 2006; *Camponotus floridanus*, Zhou et al. 2012; *Nasonia vitripennis*, Robertson et al. 2010). Branches are colored by species (Red: *Camponotus floridanus*; Light blue: *A.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

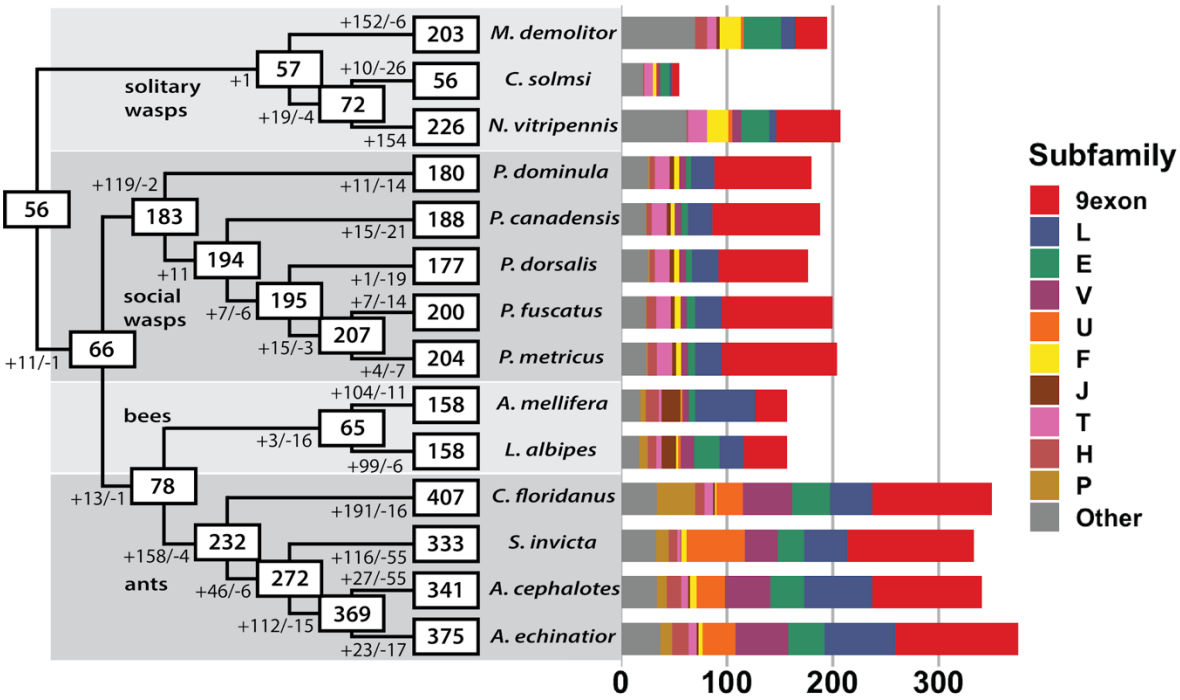

Fig. 4: Cladogram of Hymenoptera species showing estimated number of OR gene gain and loss events along branches and estimated size of ancestral and extant species OR repertoires in boxes. To the right is a bar chart showing numbers of ORs broken down by subfamily. Non-Polistine OR data are from Robertson 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.

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 (Figure 5A). 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 & Kelly 2015; Table S3). The remaining orthologous groups contain an expansion in one or more species (Figure 5B). 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 repertoires size across *Polistes* species (Figure 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 (Table S3). 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 (Figure 5B). 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.

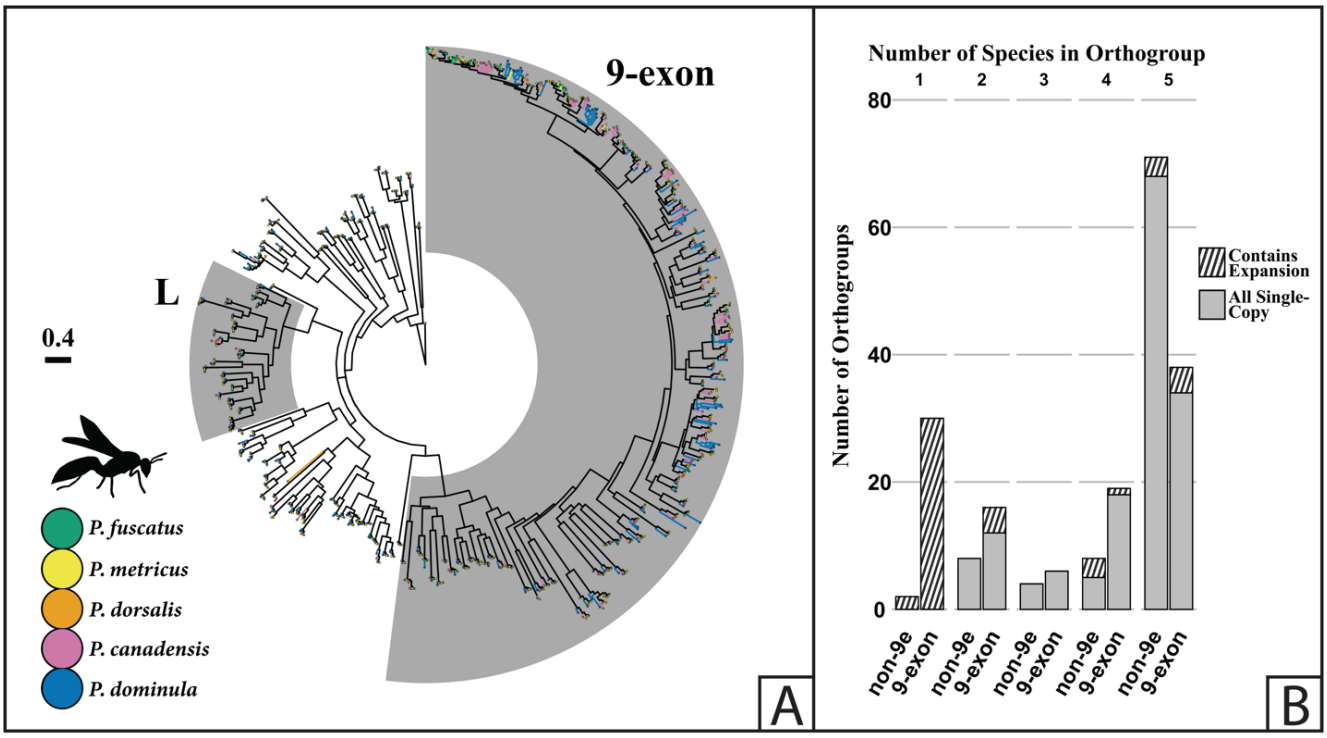

Fig. 5: (A) Maximum likelihood OR protein tree with branches colored by species (Green: *P. fuscatus*; Yellow: *P. metricus*; Orange: *P. dorsalis*; Magenta: *P. canadensis*; Blue: *P. dominula*). The L and 9-exon subfamilies are highlighted. Scale bar represents 0.4 mean substitutions per site. (B): 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 (left bar) and 9-exon orthologous groups (right bar).

272 **Microsynteny Reveals Recent Birth-and-Death Events in *Polistes* 9-exon OR Subfamily**
273 Expanded gene families often occur as tandem arrays, a genomic architecture that can contribute
274 to increased rates of gene birth and death, increasing copy number variation among species (Ohno
275 1970). Therefore, we examined how genomic organization varies between OR subfamilies in
276 *Polistes* species to generate insights into the molecular evolutionary mechanisms shaping OR
277 subfamily function. Genomic organization of ORs across *Polistes* is consistent with a model of
278 birth-and-death evolution shaping OR repertoires. As in bees, gene gain and loss at a small number
279 of loci containing tandem arrays is responsible for most copy number variation in the OR family
280 across closely-related species (Brand & Ramírez 2017). In *P. fuscatus*, 62% of ORs occur in
281 tandem arrays of 6 or more genes (Figure 6A). The frequency of tandem arrays and the tail-to-
282 head orientations of neighboring genes point to tandem duplication as the primary mechanism of
283 OR expansion, likely caused by non-allelic homologous recombination (Lynch 2007; Ramdya &
284 Benton 2010). We examined microsynteny among genes and pseudogenes in the four longest
285 tandem arrays of ORs in *Polistes* genomes (Figure 6B). There is marked decrease in OR synteny
286 among genes in orthologous 9-exon OR arrays compared to tandem arrays of L and T subfamily
287 ORs in the *Polistes* genus. The longest OR gene tandem array in *P. fuscatus* is comprised of 44
288 genes in the 9-exon subfamily on scaffold 13 (*s13*), which corresponds to homologous arrays of
289 50 genes in *P. metricus*, 25 genes in *P. dorsalis*, 33 genes in *P. canadensis*, and 29 genes in *P.*
290 *dominula*. Only 34% of *P. fuscatus* ORs in this array have orthologs across all *Polistes* species
291 sampled (Figure 6B). The second longest OR gene tandem array in *P. fuscatus* contains 24 ORs in
292 the L subfamily on scaffold 17 (*s17*), and these ORs show 1:1 orthology across *P. fuscatus*, *P.*
293 *metricus*, and *P. dorsalis*, while *P. canadensis* possesses an array of ~23 genes split between two
294 scaffolds, and *P. dominula* possesses an array of 21 ORs at this locus (Figure 6B). This tandem
295 array, widely expanded across Hymenoptera, has been expanded and conserved across *Polistes*.
296 The T subfamily, located on scaffold 8 (*s8*) of the *P. fuscatus* genome, is composed of 14 tandemly
297 arrayed genes that show 1:1 orthology across five *Polistes* (Figure 6B). Differences in the extent
298 of microsynteny among tandem arrays belonging to different OR subfamilies highlight the unique
299 evolutionary processes shaping 9-exon OR evolution in paper wasps. At the same time, the
300 extreme conservation of L and T subfamily tandem arrays across species highlights the strong
301 conservation of the OR repertoire outside of the 9-exon subfamily among *Polistes* species.

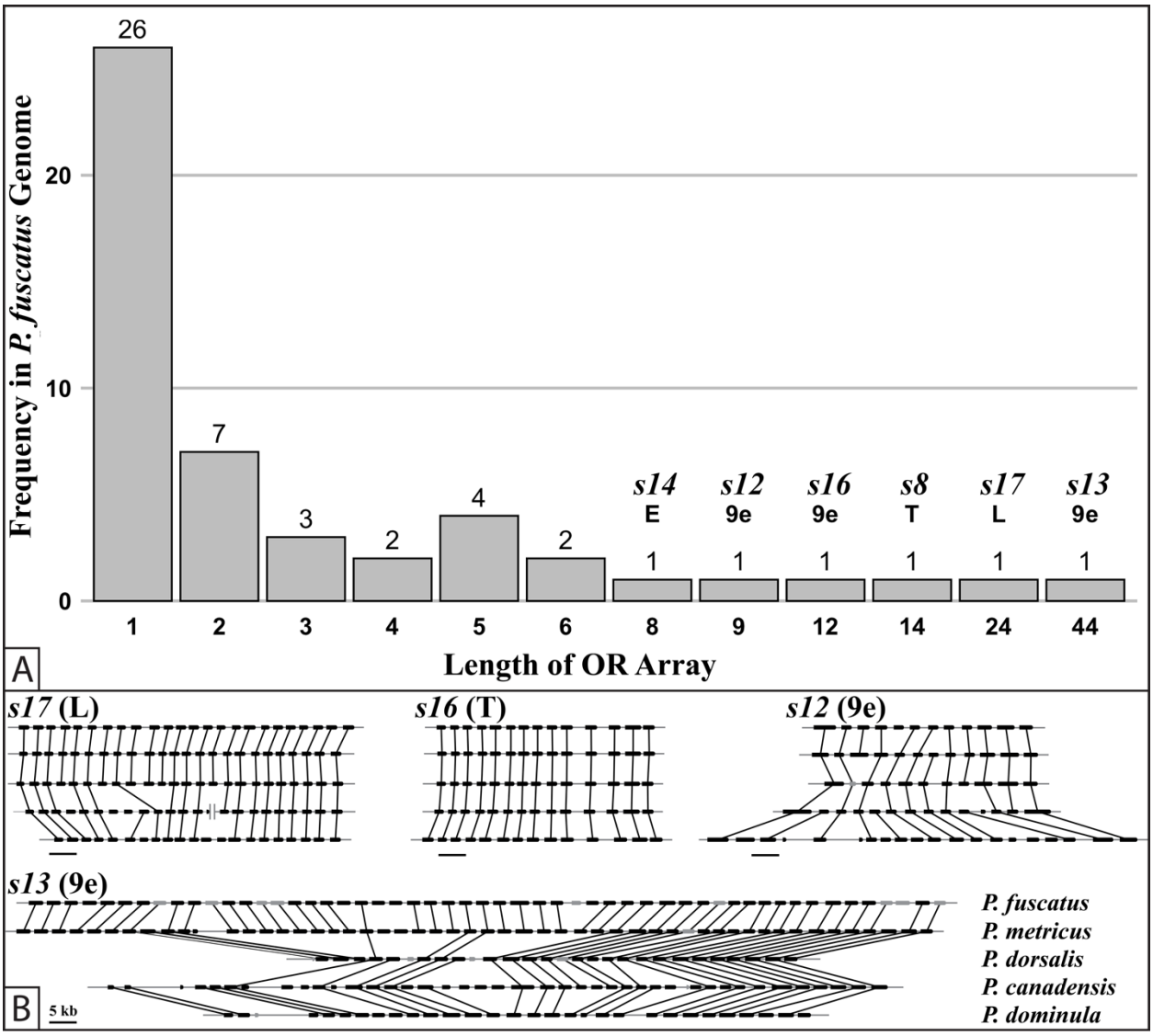

Fig. 6: (A) Frequency of OR gene singletons and tandem arrays in the *P. fuscatus* genome. 62% of ORs in *P. fuscatus* occur in tandem arrays of 6 or more genes. The longest tandem array is a 44 gene cluster on scaffold 13 (*s13*) containing 9-exon subfamily ORs. The first row of x-axis labels is the 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 (abbreviated *s#* in parentheses). Black boxes represent presumably functional genes 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. The black scale bars represent 5 kb.

314 Microsynteny analysis suggests a process of ongoing gene turnover in 9-exon arrays but stasis in
315 most other expanded subfamilies. More recent turnover should be associated with higher pairwise
316 amino acid identity between neighboring genes in an array if they are the result of recent
317 duplication events (Ohno 1970; Bohbot et al. 2007). To explore the relationship between amino
318 acid divergence and tandem array locus, we compared the mean percent amino acid identity among
319 neighboring genes within an array between the eight loci containing the longest tandem arrays of
320 ORs in the *P. fuscatus* genome using one-way ANOVA (Figure 7). Mean percent amino acid
321 identity of neighboring genes was significantly separated by OR array identity (DF = 7; F = 5.39;
322 P = 2.67e-05). Differences between particular OR tandem arrays were identified using Tukey HSD
323 post hoc tests. The mean percent amino acid identity among neighboring genes within one tandem
324 array of nine 9-exon ORs on scaffold 12 (*s12*) of the *P. fuscatus* genome is higher than in the *s13*
325 9-exon array (P Adj = 0.04586), the *s17* L array (P Adj = 0.00013), the s8 T array (P Adj =
326 0.00458), the *s16* 9-exon array (P Adj = 0.00124), and the *s19* V array (P Adj = 0.02247). The *s12*
327 9-exon OR array is composed of a larger proportion of pseudogenes (5 PSE, 10 intact gene models)
328 than the other two 9-exon arrays (*s13*: 9 PSE, 44 intact gene models; *s16*: 0 PSE, 12 intact gene
329 models). ORs in the *s12* 9-exon array lack clear orthologous relationships with ORs in species
330 other than *P. metricus*. Taken together, the high within array sequence similarity, high frequency
331 of pseudogenes, and low orthology exhibited by this array indicate that it is the result of one or
332 more recent gene duplication events since the divergence of *P. fuscatus* and *P. metricus* from the
333 other three *Polistes* species. The *s6* H subfamily array also shows higher amino acid sequence
334 identity among neighboring genes than the *s17* L subfamily array (P Adj = 0.01625) and the *s16*
335 9-exon array (P Adj = 0.04038). Increased amino acid similarity may also occur within older
336 tandem arrays as a result of gene conversion (Nagawa et al. 2002). However, we searched for gene
337 conversion using GENECONV (Sawyer 1989) and did not detect gene conversion events within
338 the *s12* 9-exon array or in the *s6* H array after Bonferroni correction. Patterns of genomic
339 organization of OR genes in *Polistes* genomes lead to the conclusion that gene gain and loss in the
340 9-exon OR subfamily is an ongoing process within this genus, in contrast to the stable and
341 conserved tandem arrays in most other OR subfamilies.

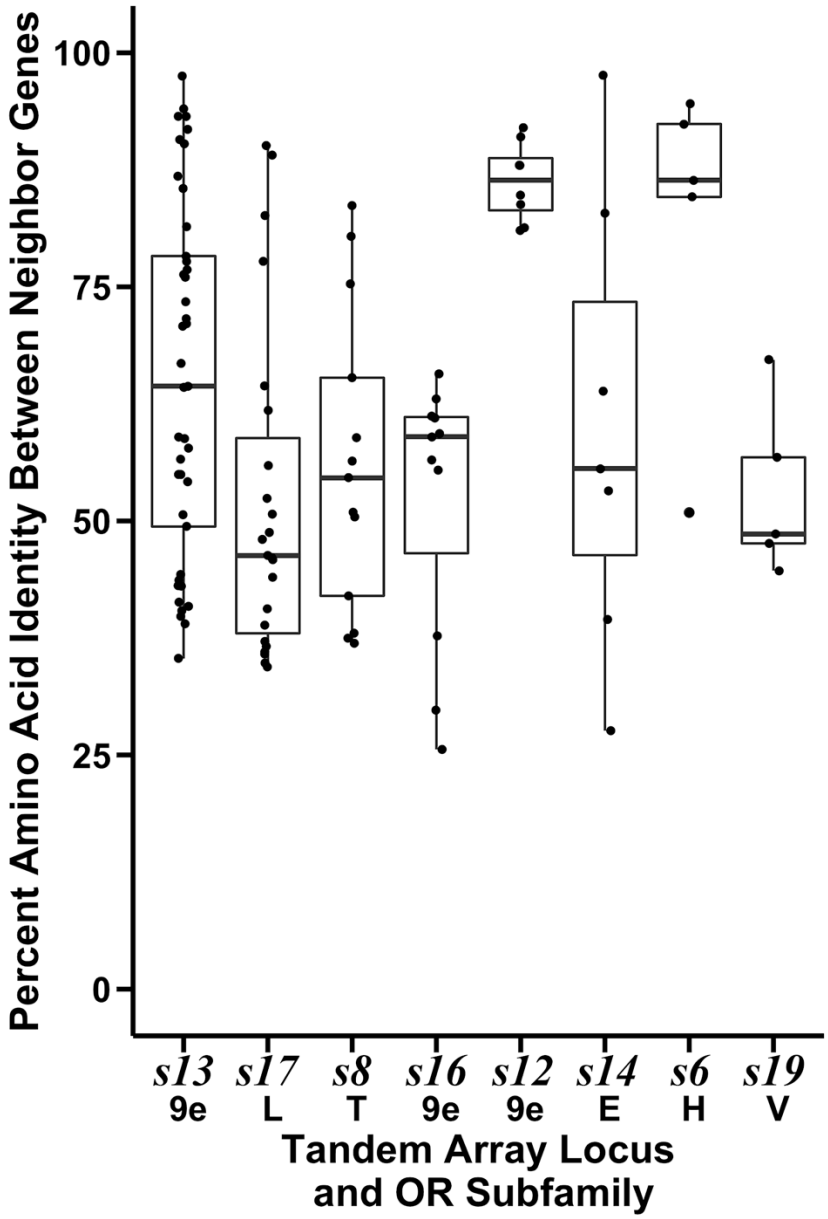

Fig. 7: Percent amino acid identity between neighboring genes at eight loci containing the longest OR gene tandem arrays in the *P. fuscatus* genome. Arrays are ordered by length in gene number, from longest (44 9-exon subfamily ORs in the *s13* tandem array) to shortest (6 H subfamily ORs in the *s6*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

356 **Positive Selection in Expanded OR Subfamilies and Accelerated Evolution of 9-exon ORs**
357 We examined patterns of OR evolution among five *Polistes* species to test the prediction that OR
358 subfamilies exhibit signatures of positive selection. We were especially interested in whether the
359 recent dynamism in 9-exon OR gene copy number has been accompanied by episodes of positive
360 selection in *Polistes*. HyPhy aBSREL (Smith MD et al. 2015; Pond et al. 2020) analyses of *Polistes*
361 OR subfamilies detected eight branches under episodic positive diversifying selection, all in OR
362 subfamilies with expansions: three branches in the 9-exon subfamily (0.33% of 918 9-exon
363 subfamily branches; Figure S4); three branches in the L subfamily (1.28% of 234 L subfamily
364 branches; Figure S5); one branch in the E subfamily (1.67% of 60 E subfamily branches; Figure
365 S6); and one branch in the H subfamily (1.54% of 65 H subfamily branches; Figure S7). This
366 supports the hypothesis that gene duplication releases duplicate genes from selective constraints,
367 allowing duplicate sequences to evolve towards other evolutionary optima (Ohno 1970). While the
368 9-exon OR subfamily is not unique among expanded OR subfamilies in its instances of episodic
369 positive selection as measured by HyPhy aBSREL, the rate of amino acid divergence is higher
370 among the 9-exon OR subfamily as a whole. The 91% mean amino acid identity among 1:1 9-exon
371 subfamily orthologs in *Polistes* is significantly lower than the 95% mean amino acid identity
372 among 1:1 orthologs in all other OR subfamilies (Figure S8; Welch Two Sample t-test: P = 7.051e-
373 05).

374 To further evaluate the patterns of nucleotide substitution driving accelerated amino acid evolution
375 of 9-exon ORs, we computed the values of d_N and d_S for pairwise alignments of 150 single copy
376 orthologs between *P. fuscatus* and *P. dorsalis* (Figure 8) using model yn00 of PAML (Yang 2007).
377 Values of d_N are significantly higher in 9-exon (mean $d_N = 0.015$) compared to other OR ortholog
378 pairs (mean $d_N = 0.006$) (Welch Two Sample t-test, P = 5.317e-07). Values of d_S are not
379 significantly elevated among 9-exon ortholog pairs compared to other OR subfamilies (mean d_S =
380 0.029 in 9-exon ORs and 0.025 in non-9exon ORs, P = 0.343). Omega values (d_N/d_S) greater than
381 1 are often considered evidence of positive selection, while $d_N/d_S = 1$ corresponds to neutral drift,
382 and $d_N/d_S < 1$ is evidence of negative selection. The omega value (d_N/d_S) for all genes is less than
383 one, suggesting negative selection. However, omega is significantly higher in 9-exon ORs than in
384 non-9-exon ORs, indicating that negative selection is weaker on 9-exon ORs (Welch Two Sample
385 t-test: mean omega = 0.32 in non-9-exon ORs and 0.644 in 9-exon ORs, P = 8.027e-05). In general,

negative selection conserves ORs shared by *P. fuscatus* and *P. dorsalis* (mean omega = 0.454), but an elevated rate of non-synonymous substitutions in 9-exon ortholog pairs imply relaxed negative selection and more drift responsible for sequence evolution in the 9-exon relative to other OR subfamilies.

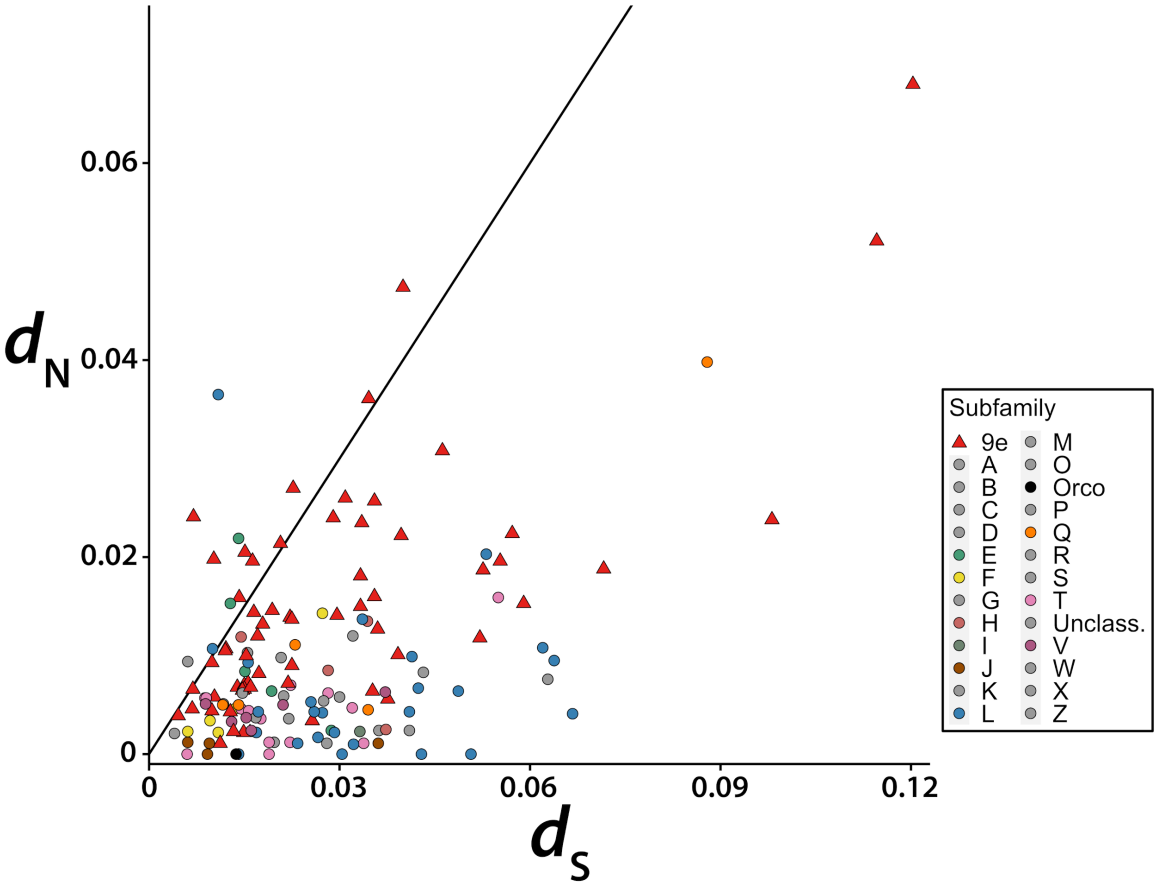

Fig. 8: The values of d_S (x-axis) and d_N (y-axis) from pairwise alignments of *P. fuscatus* and *P. dorsalis* 1:1 orthologs. Values of d_N are elevated in the 9-exon OR subfamily (data points represented by red triangles) relative to other OR subfamilies (data points represented by circles). The diagonal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

395 **DISCUSSION**

396 **Expansion of 9-exon OR Subfamily During Independent Evolution of Sociality in Wasps**

397 By carefully annotating the OR repertoires of five social wasp species spanning ~40 million years
398 of divergence in the *Polistes* genus, this study adds a higher resolution lens to our view of the
399 evolution of social insect odorant receptors. During the diversification of *Polistes,* evolutionary
400 patterns show genus-wide conservation of their ~200 ORs except for the 9-exon genes, which
401 show elevated turnover and lower sequence conservation. The 9-exon OR subfamily has
402 dramatically expanded in paper wasps, and now makes up over half of the *Polistes* OR gene set.
403 Social and ecological niches are relatively conserved within *Polistes*, though there is considerable
404 variation in social behavior and ecological niches among vespid wasps (Ross & Matthews 1991;
405 O'Neill 2001). For example, an analysis of three new, high-quality hornet genomes suggested that
406 the highly eusocial hornets have larger OR repertoires compared to the primitively eusocial
407 *Polistes* (Harrop et al. 2020). That analysis recovered less than half of the ORs reported here for
408 *Polistes*, likely due to a lack of antennal transcriptome data, suggesting that hornets may have
409 larger OR repertoires than reported. Evidence from the hornet *Vespa velutina*, including the
410 discovery of 264 antennal lobe glomeruli, indicates that the hornet OR repertoire has expanded
411 (Couto et al. 2016, 2017). Future analysis of additional genomes and antennal transcriptomes of
412 diverse social and solitary vespid wasps will allow further examination of the relationship between
413 social behavior and OR subfamily expansion.

414 **Combinatorial Coding of CHCs by 9-exon ORs Facilitates Recognition**

415 Electrophysiological deorphanization studies of 9-exon ORs in the ant *Harpegnathos saltator* offer
416 key insights into how 9-exon OR coding might relate to gene expansion. Through combinatorial
417 coding, 9-exon ORs can detect a large variety of structurally diverse CHCs. Pask et al. (2017)
418 examined 22 *H. saltator* 9-exon ORs, a subset of the 118 annotated 9-exon ORs in this species,
419 and found that 9-exon ORs were responsive to CHCs, and overlapped in their responses to multiple
420 CHC compounds. The combined responses of these 22 ORs to CHC extracts from different castes
421 were sufficient to map the CHC profiles of males, workers, and reproductive females (gamergates)
422 to separate regions of a 22-dimensional receptor space (Pask et al. 2017). This highlights the ability
423 of 9-exon ORs to facilitate social recognition by combinatorial coding. In social insect colonies,
424 CHC variation holds information at multiple levels of conspecific recognition, from inter-colony

425 nestmate recognition to within colony individual recognition (d'Ettorre & Moore 2008; Leonhardt
426 et al. 2016). Expansion of the 9-exon OR subfamily might result from selection for more
427 combinations of ORs that together can discriminate between subtle qualitative and quantitative
428 variations in CHC blends of conspecifics. Nest-specific quantitative variation in CHCs has been
429 documented across *Polistes* species (Espelie et al. 1990; Singer et al. 1992; Espelie et al. 1994;
430 Layton et al. 1994), but the molecular mechanisms underlying nestmate recognition in *Polistes* are
431 still obscure. Increased copy number of 9-exon ORs may not only expand the qualitative range of
432 compounds perceived by paper wasps, but also the quantitative olfactory space, since wasps may
433 be able to discern unique concentration differences between CHC blends as a result of the
434 combined action of 9-exon ORs with various response thresholds. Gene duplication can also
435 promote regulatory diversification (Kucharski et al. 2016; Dyson & Goodisman 2020). In *P.*
436 *metricus*, CHCs vary between castes and across stages of the colony cycle (Toth et al. 2014).
437 Regulatory subfunctionalization of duplicate ORs could be responsible for caste- and colony
438 phase-specific expression of ORs involved in detecting caste-specific and seasonally variable
439 CHCs. In addition to adaptive expansion of ORs, neutral processes contribute to OR gene birth-
440 and-death events. There may be an advantage for a large 9-exon OR gene copy number up to a
441 point, followed by random gene duplication and deletion around this optimal copy number. This
442 random genomic drift has been proposed to shape vertebrate olfactory receptor evolution and copy
443 number variation in other large multigene families (Nei 2007; but see Hayden et al. 2010).

444 **Evolution of Odorant Receptors Reflects Specific Chemosensory Ecologies of Species**

445 Social insect species differ in their level of sociality and extent of olfactory recognition abilities
446 (d'Ettorre & Moore 2008; Rehan & Toth 2015). Some social aspects of the *Polistes* colony cycle
447 vary across species. For example, the average number of cooperative foundresses varies from 1 to
448 ~6, and average sizes of mature nests may vary ~60 to ~490 cells (Reeve 1991; Sheehan et al.
449 2015; Miller SE et al. 2018). Increased 9-exon OR copy number may facilitate complex olfactory
450 recognition in species with larger colony sizes, higher cooperative nest-founding rates, and greater
451 sympatry with related species. However, expansions of 9-exon ORs are not exclusive to social
452 wasps, suggesting that the specific chemical ecology of an insect is a more influential factor
453 shaping OR evolution than level of sociality (Karpe et al. 2017). Furthermore, a meta-analysis
454 found that the complexity of CHC phenotypes does not differ between social and solitary

455 Hymenopteran species (Kather & Martin 2015). The CHC profile of *Nasonia vitripennis* includes
456 at least 52 CHC compounds, and detection of CHCs on prey items may help *Microplitis* identify
457 prey (Lewis et al. 1988; Niehuis et al. 2011). The need for solitary wasps to perceive CHCs could
458 explain why *N. vitripennis* and *M. demolitor* exhibit expansions in the 9-exon OR subfamily.

459 **Lineage-Specific Chemical Signaling in *Polistes* and Molecular Evolution of ORs**

460 Most expanded OR subfamilies are highly conserved in copy number across five *Polistes* species,
461 with the exception of the 9-exon OR subfamily. In particular, one portion of the 9-exon subfamily
462 arranged in a single tandem array (*P. fuscatus* 9e *s13*) has experienced dynamic evolution. What
463 might drive rapid gain and loss of 9-exon ORs? Divergent social chemical landscapes between
464 species may cause gene turnover as 9-exon OR evolution tracks evolutionarily labile chemical
465 signals. For example, *P. fuscatus* and *P. metricus* are closely related, and both species possess
466 CHC profiles consisting of linear and methyl-branched alkanes (Espelie et al. 1990; Espelie et al.
467 1994). However, the *P. fuscatus* CHC profile includes a higher proportion of alkenes than *P.*
468 *metricus* or *P. dominulus*, and the position of the methylated carbon of methyl-branched alkanes
469 is sometimes shifted between species (Espelie et al. 1990; Singer et al. 1992; Espelie et al. 1994;
470 Layton et al. 1994). Ant 9-exon ORs respond differently to subtle variations in CHC structure
471 (Pask et al. 2017). Between closely related *Polistes* species, structural isomers of methyl-branched
472 alkanes probably activate different ensembles of ORs.

473 If a chemical develops new behavioral relevance in a lineage, gene duplication would allow the
474 olfactory system to explore chemical space in the direction of this compound. HyPhy aBSREL
475 analyses identified eight branches in expanded OR subfamilies, including the 9-exon subfamily,
476 that have undergone positive selection during the last ~40 million years, consistent with
477 neofunctionalization or subfunctionalization of duplicated genes. Signatures of positive selection
478 on OR genes may indicate directional selection to perceive species-specific chemical signals.
479 Perception of such species-specific CHCs might be important in mate compatibility recognition.
480 In *Polistes*, mating occurs at territories defended by males and often frequented by multiple species
481 (Post & Jeanne 1983; Reed & Landolt 1990). However, the frequency of interspecific mating is
482 low, suggesting *Polistes* use vision and/or olfaction to inform their mating decisions (Miller SE et
483 al. 2019). Duplication and deletion of ORs would facilitate evolution of species-specific chemical

484 signaling systems that could contribute to reproductive isolation of sympatric species. If a chemical
485 signal is lost in a species, the corresponding ORs may become obsolete, and would be expected to
486 pseudogenize and be purged from the genome. Duplication and deletion of ORs could also lead to
487 species-specific chemical signaling in the absence of evolutionary change in chemical signals
488 (Cande et al. 2013). However, OR evolution is not strictly necessary for such a difference to evolve
489 between species, and circuit-level changes can prescribe new valence to chemical signals that are
490 shared between species and perceived by common peripheral receptors (Seeholzer et al. 2018).

491 **Conservation of Most OR Subfamilies Suggests Conserved Functions**

492 Aside from the 9-exon OR subfamily, gene expansions have occurred in subfamilies L, T, H, E,
493 and V (Figure 4). A larger variety of ORs relaying information through ORNs to a larger number
494 of antennal lobe glomeruli will increase sensory acuity in any olfactory discrimination task, social
495 or otherwise. An ancient locus of tandemly duplicated L subfamily ORs observed across social
496 insects has expanded in *Polistes*, although to a lesser extent than in other social insects (~50 L
497 subfamily ORs in honeybee and ants, 25 L subfamily ORs in a tandem array on *P. fuscatus* scaffold
498 17). Odorant receptors in the L subfamily are thought to detect queen pheromone components and
499 fatty acids in bees as well as CHCs in ants (Wanner et al. 2007; Karpe et al. 2016; Pask et al. 2017).
500 The T subfamily has expanded to a greater degree in *P. fuscatus* (14 genes) than in ants (~7 genes)
501 and the honeybee (2 genes), but no ORs in this clade have been functionally characterized. The *P.*
502 *fuscatus* genome encodes nine H subfamily ORs, which are putative floral odorant detectors in
503 bees, and which also respond to CHCs and other general odorants in ants (Claudianos et al. 2014;
504 Slone et al. 2017). Fatty acids and volatile organic compounds are produced by flowers that wasps
505 rely on as a source of carbohydrates (Raguso 2008). Expansions in several OR subfamilies may
506 increase olfactory discrimination of chemicals with diverse behavioral relevance. *Polistes* species
507 are distributed globally in temperate and tropical regions, occupying similar social and ecological
508 niches as generalist predators and floral foragers that form primitively eusocial societies (Reeve
509 1991; Richter 2000). A conserved set of ORs may perform common functions in conserved
510 behaviors across paper wasp species. High levels of OR conservation are also consistent with a
511 specialist molecular function of an OR in a dedicated channel of olfaction. Patterns of molecular
512 evolution suggest conserved behavioral and molecular functions of most non-9-exon OR
513 subfamilies in *Polistes*.

514 **Distinct Patterns of OR Evolution Within the Same Genome**

515 The differences between the conserved OR repertoires in *Drosophila* and the more dynamic
516 evolution of vertebrate OR gene families have given rise to speculation about the relationship
517 between OR function and evolution (Nozawa & Nei 2007; Andersson et al. 2015). There is
518 prevalent negative selection conserving odorant receptors across *Drosophila* species, and the
519 majority of *D. melanogaster* ORs form simple orthologous relationships across the genus (Clark
520 et al. 2007; Guo & Kim 2007; McBride & Arguello 2007; Nozawa & Nei 2007; Sánchez-Gracia
521 et al. 2009; Mansourian & Stensmyr 2015). In paper wasps we report both highly conserved OR
522 expansions similar to those seen in *Drosophila* as well as elevated gene turnover and drift among
523 the 9-exon ORs, reminiscent of a more vertebrate-like evolutionary pattern. If the highly dynamic
524 clades of 9-exon ORs of social wasps are involved in more combinatorial coding compared to
525 other more conserved 9-exon or non-9-exon ORs, that would indicate a link between molecular
526 evolution of odorant receptors and neural coding. Further investigations into the relative tuning of
527 9-exon as well as more conserved ORs in social wasps and other social insects provide a promising
528 research direction to investigate the links between molecular evolutionary patterns, odorant
529 receptor tuning, and neural coding.

530 ## MATERIALS & METHODS

531 **Antennal Lobe Imaging**

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

535 **Gene Annotation**

536 The *P. fuscatus*, *P. dorsalis*, and *P. metricus* genomes were assembled and automatedly annotated
537 as described in Miller SE et al. (2020). The *P. canadensis* and *P. dominula* genomes and
538 annotations were accessed through NCBI (Patalano et al. 2015; Standage et al. 2016). Coding
539 regions of ORs were identified by using TBLASTN (Altschul et al. 1997) with a sample of OR
540 proteins from 19 insect species used as query sequences. Genomes were queried iteratively with
541 TBLASTN, adding newly annotated *Polistes* ORs to the query file, until no new OR coding regions
542 were identified. To guide annotation of exon-intron boundaries, antennal mRNA from *P. fuscatus*
543 males and females (gynes) was mapped to *P. fuscatus*, *P. metricus*, and *P. dorsalis* genomes using
544 STAR (Dobin et al. 2013) and assembled into transcripts using Trinity (Table S2; Haas et al. 2013).
545 Predicted transcripts were aligned to genomes using BLAT. Uncertain gene models in *P. metricus*,
546 *P. dorsalis*, *P. canadensis*, and *P. dominula* were aligned to their orthologs in *P. fuscatus* using
547 Muscle version 3.8.425 with maximum 4 iterations (Edgar 2004), and gene models were manually
548 adjusted. All annotation evidence was imported into Geneious v11.1.5 genome browser for manual
549 annotation. The majority of apparently functional ORs that were not detected by the automated
550 annotation and required extensive manual curation were 9-exon subfamily receptors (e.g. Figure
551 S9). Gene models were called pseudogenes if they exhibited frame-shift mutation, premature stop
552 codons, or unacceptable 5' donor or 3' receptor splice sites. Transmembrane helices of all
553 putatively functional ORs (>300 amino acids) were predicted using TMHMM version 2.0c
554 (Sonnhammer et al. 1998) and Phobius version 1.01 (Käll et al. 2004). Throughout the main text,
555 "putatively functional ORs" are OR proteins at least 300 amino acids in length. Details of mRNA
556 library preparation, sequencing, read mapping, and manual gene annotation are included as
557 Supplementary Materials & Methods.

558 **Phylogenetic Reconstruction**

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

565 **Genomic Organization**

566 OR genes and pseudogenes were considered to be in a tandem array if they were uninterrupted by
567 non-OR genes and were within 5kb of each other. The lengths of OR arrays correspond to the
568 number of putatively functional ORs and exclude the pseudogenes contained within the array. The
569 pairwise percent amino acid identity between neighbors in an array was calculated using only
570 putatively functional ORs that neighbored another putatively functional OR within 5kb. See
571 Supplementary Materials & Methods for details.

572 **Sequence Analyses**

573 All putatively functional *Polistes* ORs greater than 350 amino acids in length were used in analyses
574 of episodic diversifying positive selection and gene conversion within OR subfamilies using
575 aBSREL in HyPhy version 2.5.15 (Smith MD et al. 2015; Pond et al. 2020) and GENECONV
576 (Sawyer 1989). Values of pairwise d_N/d_S for orthologs shared by *P. fuscatus* and *P. dorsalis* were
577 calculated using PAML version 4.9 (Yang 2007) program yn00 with the Yang & Nielsen (2000)
578 method. See Supplementary Materials & Methods for details.

579 **Data Availability**

580 The genome assemblies analyzed in this article are available on Genbank (see Table S1). Gene
581 models, amino acid sequences, and nucleotide sequences underlying this article, as well as
582 alignments analyzed in selection analyses, are available in its online Supplementary Material.

583 **SUPPLEMENTARY MATERIAL**

584 Supplementary Materials & Methods, Supplementary Tables S1-S3, and Supplementary Figures
585 S1-S9 are included in a PDF file available in the online Supplementary Material. Meta-data
586 associated with *Polistes* ORs, including subfamily membership, transmembrane domain
587 prediction, and sequence length are included in an excel file available in the online
588 Supplementary Material.

589 **ACKNOWLEDGEMENTS**

590 This work was supported by National Science Foundation [Graduate Research Fellowship
591 Program grant number DGE-1650441 to A.W.L., CAREER grant number DEB-1750394 to
592 M.J.S.], National Institutes of Health [grant numbers DP2-GM128202 to M.J.S., S10OD018516
593 to the Cornell University Biotechnology Resource Center], and New York State Stem Cell Science
594 [grant number CO29155 to Cornell BRC]. We thank Qi Sun and the Computational Biology
595 Service Unit of the Cornell Life Sciences Core Laboratories Center for making software available
596 on BioHPC and for providing helpful advice regarding gene annotation. We also thank Brook
597 Luers and the Cornell Statistical Consulting Unit for statistics consultation.

598 **AUTHOR CONTRIBUTIONS**

599 AWL and MJS conceptualized the study. AWL sequenced mRNA, manually annotated ORs,
600 conducted all analyses, and wrote the paper. CMJ dissected and imaged antennal lobes, and
601 wrote the Supplementary Materials & Methods section on Antennal Lobe Imaging. AWL and
602 MFF labeled antennal lobe glomeruli. SEM and MJS sequenced and assembled the genomes. All
603 authors helped edit the final manuscript.

604 **REFERENCES**

605 Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST
606 and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–
607 3402.

608 Andersson MN, Löfstedt C, Newcomb RD. 2015. Insect olfaction and the evolution of receptor tuning.
609 *Front Ecol Evol.* 3:53.

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

612 Benton R. 2015. Multigene family evolution: perspectives from insect chemoreceptors. *Trends Ecol Evol.*
613 30:590-600.

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

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

617 Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data.
618 *Bioinformatics.* 30:2114–2120.

619 Brand P, Ramírez SR, Leese F, Quezada-Euan JG, Tollrian R, Eltz T. 2015. Rapid evolution of
620 chemosensory receptor genes in a pair of sibling species of orchid bees (Apidae: Euglossini). *BMC Evol*
621 *Biol.* 15:176.

622 Brand P, Ramirez SR. 2017. The evolutionary dynamics of the odorant receptor gene family in
623 corbiculate bees. *Genome Biol Evol.* 9:2023-2036.

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

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

628 Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated alignment
629 trimming in large-scale phylogenetic analyses. *Bioinformatics.* 25:1972–1973.

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

632 Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA et al. 2007. Evolution of genes
633 and genomes on the *Drosophila* phylogeny. *Nature.* 450:203-218.

634 Claudianos C, Lim J, Young M, Yan S, Cristino AS, Newcomb RD, Gunasekaran N, Reinhard J. 2014.
635 Odor memories regulate olfactory receptor expression in the sensory periphery. *European J Neurosci.*
636 39:1642-1654.

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

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

641 Dani FR, Jones GR, Destri S, Spence SH, Turillazzi S. 2001. Deciphering the recognition signature
642 within the cuticular chemical profile of paper wasps. *Anim Behav.* 62:165-171.

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

645 Dani FR, Turillazzi S. 2018. Chemical communication and reproduction partitioning in social wasps. *J*
646 *Chem Ecol.* 44:796-804.

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

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

652 Dobin A, Davis CA, Schlesinger F, Drnkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013.
653 STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2:15-21.

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

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

658 Ebrahim SAM, Dweck HKM, Stökl J, Hofferberth JE, Trona F, Weniger K, Rybak J, Seki Y, Stensmyr
659 MC, Sachse S, Hansson BS, Knaden M. 2015. *Drosophila* avoids parasitoids by sensing their
660 semiochemicals via a dedicated olfactory circuit. *PLoS Biol*. 13:e1002318.

661 Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput.
662 *Nucleic Acids Res.* 32:1792–1797.

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

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

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

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

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

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

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

677 Fishilevich E, Vosshall LB. 2005. Genetic and functional subdivision of the *Drosophila* antennal lobe.
678 *Curr Biol.* 15:1548-1553.

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

681 Gamboa GJ, Reeve HK, Pfennig DW. 1986. The evolution and ontogeny of nestmate recognition in social
682 wasps. *Ann Rev Entomol.* 31:431-454.

683 Gamboa GJ, Grudzien TA, Espelie KE, Bura EA. 1996. Kin recognition in social wasps: combining
684 chemical and behavioural evidence. *Anim Behav.* 51:625-629.

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

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

690 Guo S, Kim J. 2007. Molecular evolution of *Drosophila* odorant receptor genes. *Mol Biol Evol.* 24:1198-
691 1207.

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

695 Hallem EA, Ho MG, Carlson JR. 2004. The molecular basis of odor coding in the *Drosophila* Antenna.
696 *Cell.* 117:965-979.

697 Hallem EA, Carlson JR. 2006. Coding of odors by a receptor repertoire. *Cell.* 125:143-160.

698 Harrop TWR, Guhlin J, McLaughlin GM, Perminia E, Stockwell P, Gilligan J, Le Lec MF, Gruber
699 MAM, Quinn O, Lovegrove M, et al. 2020. High-quality assemblies for three invasive social wasps from
700 the Vespula genus. *G3-Genes Genom Genet. Early online August 28, 2020*;
701 https://doi.org/10.1534/g3.120.401579

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

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

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

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

710 Käll L, Krogh A, Sonnhammer ELL. 2004. A combined transmembrane topology and signal peptide
711 prediction method. *J Mol Biol.* 338:1027-1036.

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

715 Karpe SD, Jain R, Brockmann A, Sowdhamini R. 2016. Identification of complete repertoire of *Apis*
716 *florea* odorant receptors reveals complex orthologous relationships with *Apis mellifera*. *Genome Biol*
717 *Evol.* 8:2879-2895.

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

721 Karpe SD, Tiwari V, Sowdhamini R. 2020. InsectOR - webserver for sensitive identification of insect
722 olfactory receptor genes from non-model genomes. *bioRxiv*, doi:
723 https://doi.org/10.1101/2020.04.29.067470, last accessed September 4, 2020.

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

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

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

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

733 Lavine BK, Morel L, Vander Meer RK, Gunderson RW, Han JH, Bonanno A, Stine A. 1990. Pattern
734 recognition studies in chemical communication: nestmate recognition in *Camponotus floridanus*.
735 *Chemometr Intell Lab.* 9:107-114.

736 Leonhardt SD, Menzel F, Nehring V, Schmitt T. 2016. Ecology and evolution of communication in social
737 insects. *Cell.* 164:1277-1287.

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

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

742 Malnic B, Hirono J, Sato T, Buck LB. 1999. Combinatorial receptor codes for odors. *Cell.* 96:713-723.

743 Mansourian S, Stensmyr MC. 2015. The chemical ecology of the fly. *Current Opinion in Neurobiology*
744 34:95-102.

745 Mathew D, Martelli C, Kelley-Swift E, Brusalis C, Gershow M, Samuel ADT, Emonet T, Carlson JR.
746 2013. Functional diversity among sensory receptors in a *Drosophila* olfactory circuit. *Proc Natl Acad Sci*
747 *USA.* 110:E2134-E2143.

748 McBride CS, Arguello JR. 2007. Five *Drosophila* genomes reveal nonneutral evolution and the signature
749 of host specialization in the chemoreceptor superfamily. *Genetics.* 177:1395-1416.

750 McKenzie SK, Fetter-Pruneda I, Ruta V, Kronauer DJC. 2016. Transcriptomics and neuroanatomy of the
751 clonal raider ant implicate an expanded clade of odorant receptors in chemical communication. *Proc Natl*
752 *Acad Sci USA.* 49:14091-14096.

753 McKenzie SK, Kronauer DJC. 2018. The genomic architecture and molecular evolution of ant odorant
754 receptors. *Genome Res.* 28:1757-1765.

755 Miller CH, Campbell P, Sheehan MJ. 2020. Distinct evolutionary trajectories of V1R clades across mouse
756 species. *BMC Evol Biol.* 20:99.

757 Miller SE, Bluher SE, Bell E, Cini A, Da Silva RC, De Souza AR, Gandia KM, Jandt J, Loope K, Prato A
758 et al. 2018. WASPnest: a worldwide assessment of social Polistine nesting behavior. *Ecology.* 99:2405.

759 Miller SE, Legan AW, Flores ZA, Ng HY, Sheehan MJ. 2019. Strong, but incomplete, mate choice
760 discrimination between two closely related species of paper wasp. *Biol J Linn Soc.* 126:614-622.

761 Miller SE, Legan AW, Henshaw M, Ostevik KL, Samuk K, Uy FMK, Sheehan MJ. 2020. Evolutionary
762 dynamics of recent selection for enhanced social cognition. *Proc Natl Acad Sci USA.* 117:3045-3052.

763 Münch D, Galizia CG. 2016. DoOR 2.0 - comprehensive mapping of *Drosophila melanogaster* odorant
764 responses. *Sci Rep.* 6:21841.

765 Nagawa F, Yoshihara S-i, Tsuboi A, Serizawa S, Itoh K, Sakano H. 2002. Genomic analysis of the
766 murine odorant receptor *MOR28* cluster: a possible role of gene conversion in maintaining the olfactory
767 map. *Gene.* 292:73-80.

768 Nei M. 2007. The new mutation theory of phenotypic evolution. *Proc Natl Acad Sci USA.* 104:12235-
769 12242.

770 Nei M. 2013. Mutation-Driven Evolution. Oxford (UK): Oxford University Press.

771 Niehuis O, Büllesbach J, Judson AK, Schmitt T, Gadau J. 2011. Genetics of cuticular hydrocarbon
772 differences between males of the parasitoid wasps *Nasonia giraulti* and *Nasonia vitripennis*. *Heredity.*
773 107:61-70.

774 Nozawa M, Nei M. 2007. Evolutionary dynamics of olfactory receptor genes in *Drosophila* species. *Proc*
775 *Natl Acad Sci USA.* 104:7122-7127.

776 Ohno S. 1970. Evolution by gene duplication. Berlin: Springer-Verlag.

777 Oi CA, Oliveira RC, van Zweden JS, Mateus S, Millar JG, Nascimento FS, Wenseleers T. 2019. Do
778 primitively eusocial wasps use queen pheromones to regulate reproduction? A case study of the paper
779 wasp *Polistes satan*. *Front Ecol Evol.* 7:199.

780 O'Neill KM. 2001. Solitary wasps: behavior and natural history. Ithaca (NY): Cornell University Press.

781 Pask GM, Slone JD, Millar JG, Das P, Moreira JA, Zhou X, Bello J, Berger SL, Bonasio R, Desplan C, et
782 al. 2017. Specialized odorant receptors in social insects that detect cuticular hydrocarbon cues and
783 candidate pheromones. *Nat Commun.* 8:297.

784 Patalano S, Vlasova A, Wyatt C, Ewels P, Camara F, Ferreira PG, Asher CL, Jurkowski TP, Segonds-
785 Pichon A, Bachman M, et al. 2015. Molecular signatures of plastic phenotypes in two eusocial insect
786 species with simple societies. *Proc Natl Acad Sci USA.* 112:13970–13975.

787 Peters RS, Krogmann L, Mayer C, Donath A, Gunkel S, Meusemann K, Kozlov A, Podsiadlowski L,
788 Petersen M, Lanfear R, et al. 2017. Evolutionary History of the Hymenoptera. *Curr Biol.* 27:1013-1018.

789 Pond SLK, Poon AFY, Velazquez R, Weaver S, Hepler NL, Murrell B, Shank SD, Magalis BR, Bouvier
790 D, Nekrutenko A, et al. 2020. HyPhy 2.5 – A customizable platform for evolutionary hypothesis testing
791 using phylogenies. *Mol Biol Evol.* 37:295-299.

792 Post DC, Jeanne RL. 1983. Male reproductive behavior of the social wasp *Polistes fuscatus*
793 (Hymenoptera: Vespidae). *Z Tierpsychol.* 62:157-171.

794 Post DC, Jeanne RL. 1984. Recognition of conspecifics and sex by territorial males of the social wasp
795 *Polistes fuscatus* (Hymenoptera: Vespidae). *Behaviour.* 91:78-92.

796 Raguso RA. 2008. Wake up and smell the roses: the ecology and evolution of floral scent. *Annu Rev Ecol*
797 *Evol Syst.* 39:549-569.

798 Ramdya P, Benton R. 2010. Evolving olfactory systems on the fly. *Trends Genet.* 26:307-316.

799 Reed HC, Landolt PJ. 1990. Sex attraction in paper wasp, *Polistes exclamans* Viereck (Hymenoptera:
800 Vespidae), in a wind tunnel. *J Chem Ecol.* 6(4):1277-1287.

801 Reeve HK. 1991. Polistes. In: Ross KG, Matthews RD, editors. The social biology of wasps. Ithaca (NY):
802 Cornell University Press. p. 99-148.

803 Rehan SM, Toth AL. 2015. Climbing the social ladder: the molecular evolution of sociality. *Trends Ecol*
804 *Evol.* 30:426-433.

805 Richter MR. 2000. Social wasp (Hymenoptera: Vespidae) foraging behavior. *Annu Rev Entomol*. 45:121-
806 150.

807 Robertson HM, Wanner KW. 2006. The chemoreceptor superfamily in the honeybee *Apis mellifera*:
808 expansion of the odorant, but not gustatory, receptor family. *Genome Res.* 16:1395-1403.

809 Robertson HM, Gadau J, Wanner KW. 2010. The insect chemoreceptor superfamily of the parasitoid
810 jewel wasp *Nasonia vitripennis*. *Insect Mol Biol.* 19:121-136.

811 Ross KG, Matthews RW. 1991. The social biology of wasps. Ithaca (NY): Cornell University Press.

812 Sadd BM, Barribeau SM, Bloch G, de Graaf DC, Dearden P, Elsik CG, Gadau J, Grimmelikhuijzen CJP,
813 Hasselmann M, Lozier JD, et al. 2015. The genomes of two key bumblebee species with primitive
814 eusocial organization. *Genome Biol.* 16:76

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

817 Sawyer SA. 1989. Statistical tests for detecting gene conversion. *Mol Biol Evol.* 6:526-538.

818 Seeholzer LF, Seppo M, Stern DL, Ruta V. 2018. Evolution of a central neural circuit underlies
819 *Drosophila* mate preferences. *Nature.* 559:564-569.

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

823 Singer TL, Camann MA, Espelie KE. 1992. Discriminant analysis of cuticular hydrocarbons of social
824 wasp *Polistes exclamans* Viereck and surface hydrocarbons of its nest paper and pedicel. *J Chem Ecol.*
825 18:785-797.

826 Singer TL. 1998. Roles of hydrocarbons in the recognition systems of insects. *Amer Zool.* 38:394-405.

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

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

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

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

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

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

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

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

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

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

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

857 Sturgis SJ, Gordon DM. 2012. Nestmate recognition in ants (Hymenoptera: Formicidae): a review.
858 *Myrmecol News.* 16:101-110.

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

862 Touhara K, Vosshall L. 2009. Sensing odorants and pheromones with chemosensory receptors. *Annu Rev*
863 *Physiol.* 71:307-332.

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

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

868 van Zweden JS, d'Ettorre P. 2010. Nestmate recognition in social insects and the role of hydrocarbons. In:
869 Blomquist GJ, Bagnères A-G, editors. Insect Hydrocarbons. Cambridge: Cambridge University Press. p.
870 222-243.

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

873 Wanner KW, Nichols AS, Walden KKO, Brockmann A, Luetje CW, Robertson HM. 2007. A honey bee
874 odorant receptor for the queen substance 9-oxo-2-decenoic acid. *Proc Natl Acad Sci USA.* 104:14383-
875 14388.

876 Wicher D. 2015. Olfactory signaling in insects. In: Glatz R, editor. Molecular Basis of Olfaction.
877 Elsevier. p. 37-54.

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

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

883 Yang Z, Nielsen R. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic
884 evolutionary models. *Mol Biol Evol.* 17:32-43.

885 Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 24:1586-1591.

886 Yapici N, Zimmer M, Domingos AI. 2014. Cellular and molecular basis of decision-making. *EMBO Rep.*
887 15:1023-1035.

888 Zhou X, Slone JD, Rokas A, Berger SL, Liebig J, Ray A, Reinberg D, Zwiebel LJ. 2012. Phylogenetic
889 and Transcriptomic Analysis of Chemosensory Receptors in a Pair of Divergent Ant Species Reveals Sex-
890 Specific Signatures of Odor Coding. *PLoS Genet.* 8, e1002930.

891 Zhou X, Rokas A, Berger SL, Liebig J, Ray A, Zwiebel LJ. 2015. Chemoreceptor Evolution in
892 Hymenoptera and its Implications for the Evolution of Eusociality. *Genome Biol Evol.* 7:2407-2416.

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