MOLECULAR AND CHEMICAL BASIS OF SOCIAL OLFACTION IN *POLISTES* PAPER WASPS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Andrew Wesley Legan

August 2022

Art by Vibhu Krishna
(Grayscale photo of canvas painting) www.vibhukrishna.com

MOLECULAR AND CHEMICAL BASIS OF SOCIAL OLFACTION IN *POLISTES* PAPER WASPS

Andrew Wesley Legan, Ph. D.

Cornell University 2022

During my graduate studies, I focused on the northern paper wasp *Polistes fuscatus*, a classic model system in the study of animal recognition behavior. *P. fuscatus* wasps use cuticular chemicals as a source of information to guide multiple behaviors, such as mate choice and nestmate recognition. I investigated phenotypic variation in *P. fuscatus* cuticular hydrocarbons (CHCs) using solid-phase microextraction and GC-FID. Chapter 1 presents the CHCs found on *P. fuscatus*, informed by GC-MS results generously shared by Dr. Kevin Loope, and identifies nest-specific chemical signatures that could mediate nestmate recognition. Chapter 2 presents comparative genomics analyses of the *Polistes* odorant receptor (OR) gene family. I argue that the OR gene family and 9-exon subfamily expanded in gene copy number during the evolution of sociality in wasps, independent of ants. Patterns of molecular evolution of the 9-exon OR subfamily are investigated, and the relevance of these evolutionary patterns to the perception of complex chemical signatures is discussed. Chapter 3 reports sex-biased antennal gene expression patterns in *P. fuscatus*, identifying candidate sex pheromone receptors and documenting expression of 9-exon ORs in males, deviating from patterns observed in ants. In Chapter 4, I analyze short-term changes in body posture of *P. fuscatus* foundresses in response to experimentally introduced conspecific intruders in the field and discuss how persistent vigilance may relate to plasticity in nestmate recognition behavior.

BIOGRAPHICAL SKETCH

Andrew Wesley Legan was born in Charleston, SC in 1992 and raised in Dyersburg, TN. He graduated from Dyersburg City Schools in 2011. He then got his B.A. in Ecology, Evolution, and Organismal Biology at Vanderbilt University in Nashville, TN in 2015, where he was advised by Drs. Patrick Abbot, Malu Jorge, and Ken Catania. He began graduate school in the Department of Neurobiology and Behavior at Cornell University in 2016, where he was advised by Dr. Michael Sheehan, Dr. Robert Raguso, Dr. Nilay Yapici, and Dr. Kern Reeve. In July 2022, he began work as an ORISE Postdoctoral Fellow with Dr. Ken Callicott at the USDA Arid-Land Agricultural Research Center Pest Management and Biocontrol Research Unit in Tucson, AZ.

Dedicated to my family

ACKNOWLEDGEMENTS

Thank you to the funding sources who have made the research I've done during my dissertation work possible. These funding sources include the National Science Foundation [Graduate Research Fellowship Program grant number DGE-1650441 to Andrew Legan, CAREER grant number DEB-1750394 to Michael Sheehan], National Institutes of Health [grant numbers DP2-GM128202 to MJS, S10OD018516 to the Cornell University Biotechnology Resource Center], North American Section of the International Union for the Study of Social Insects Robert L. and Louise B. Jeanne Social Wasp Research Grant, National Geographic Society, Cornell University Andrew W. Mellon Student Research Grant, and New York State Stem Cell Science [grant number CO29155 to Cornell BRC].

I am overwhelmed when I think about the generous sharing of time and resources that made it possible for me to carry out the research leading to this doctoral dissertation. Here I will acknowledge some of the people and groups of people who made this dissertation possible. First, I would like to thank my doctoral advisor, Dr. Mike Sheehan. Mike has been an ideal doctoral advisor who enthusiastically supported my research efforts and always gave me feedback regardless of how close to the deadline I sent him documents for comments. I also thank my Special Committee members, past and present: Dr. Robert Raguso, Dr. Nilay Yapici, Dr. Kern Reeve, and Dr. Thomas Seeley. I am indebted to the past and present members of the Laboratory for Animal Social Evolution and Recognition (LASER) and the Dept. of Neurobiology and Behavior for all their support.

I have many people to thank for their help with components of individual manuscripts (chapters). I thank Matthieu Fuchs and Arthur Chen for their assistance with laboratory research. I thank Dr. Christiane Linster for helping me find paper wasps nests needed for research experiments. To help with Chapter 1, GC-MS data for three *Polistes fuscatus* wasps was generously provided by Dr. Kevin Loope and Dr. Patrizia d'Ettorre. Ethan Bass shared his useful R package, chromatographR, as well as R scripts for linear interpolation of retention times to generate retention indices, and Callum Kingwell helped me to troubleshoot SPME sampling methods. I especially thank Dr. Callum Kingwell, Dr. Geoffrey Broadhead, and Dr. Rob Raguso for all their time spent teaching me how to operate and interpret results generated by the GC-FID. The Cornell Liddell Field Station made this work possible, and Dr. Tom Seeley's mentorship at Liddell was invaluable. For help with Chapter 2, I 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. I also thank Dr. Brook Luers and the Cornell Statistical Consulting Unit, who provided statistics consultation, as well as Matt Einhorn and Judah Sklan for help troubleshooting scripts and Dr. Caitlin Miller for feedback on the manuscript corresponding to Chapter 2. While learning how to navigate bioinformatics tools to study odorant receptors, Dr. Caitlin Miller and Dr. Angela Freeman were extremely helpful and always willing to talk about challenges and possible solutions. I was fortunate to have the opportunity to learn from the talented postdocs working in the LASER lab. Dr. Sara Miller and Dr. Chris Jernigan deserve special thanks for their time spent teaching me about insect genomics and

insect olfactory behavior. I thank the attendees of NBB Lunch Bunch for all their helpful feedback on my research projects, especially Dr. Kerry Shaw, Dr. Mike Webster, Dr. Rob Raguso, and Dr. Kern Reeve. I also thank the Dept. of Neurobiology and Behavior administration and staff, past and present, who were always helpful and supported me in countless ways.

I am grateful for the support of my parents, Mary and Mark, who on several wasp collecting trips to TN helped me fill the dining room table with cups full of wasps. I thank my supportive partner, Serena Peterson, who helped me find *Mordvilkoja vagabunda* aphid galls around Ithaca for a side research project. Many others are indirectly responsible in one way or another for the completion of this dissertation and, for fear of leaving someone out, I will extend my thanks to them individually during the coming months as I find time to reflect more intentionally.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

NEST-SPECIFIC SIGNATURES IN THE CUTICULAR HYDROCARBONS OF *POLISTES FUSCATUS* PAPER WASPS

Andrew W. Legan, Michael J. Sheehan

Abstract

Cuticular hydrocarbons (CHCs) are important signaling molecules in insects. In social insects, including paper wasps, nest-specific patterns of CHC variation are perceived by individuals on the nest, mediating nestmate recognition behavior. This study analyzed cuticular hydrocarbons of 189 female northern paper wasps, *Polistes fuscatus*, and identified nest-specific patterns of CHC profile variation between nests in the wild. Our chromatography, informed by results from GC-MS generously provided by Dr. Kevin Loope and Dr. Patrizia d'Ettorre, uncovered 96 cuticular hydrocarbon peaks, including 76 peaks previously undocumented in this species. We investigated nest-specificity of peaks representing different chemical classes of the CHC profile, which includes 11 linear alkanes, 3 linear alkenes, 3 methyl-branched alkenes, and 27 mono-, 22 di-, 14 tri-, and 4 tetra-methyl branched alkanes. Consistent with previously published work, we identified methylalkanes and some linear alkanes to be especially diagnostic of nest affiliation in *P. fuscatus*. This study extends previous work that identified mono- and di-methyl alkanes to also include tri- and tetra-methyl alkanes in the list of candidate recognition signature compounds. Our results support the hypothesis that the nestmate recognition signature of paper

wasps is a complex mixture including primarily methylalkanes, providing predictions for behavioral assays testing salience of compounds in nestmate recognition.

Background

Insect social groups usually consist of closely related individuals with shared inclusive fitness interests, which is why unrelated individuals from outside of the group are often aggressively rejected (Gadagkar 1985; Arnold et al. 1996; Breed et al. 2007). This nestmate recognition system ensures that the fecundity of the nest is uncompromised by conflicting interests of conspecific or heterospecific non-nestmates. The process is hypothesized to occur by phenotype-matching: a recognition mechanism based on comparing the phenotypic label of an individual to a recognition template or templates stored in memory (Lacy and Sherman 1983; Reeve 1989; Tsutsui 2004; Neupert et al. 2018). If the template-label dissimilarity is below a threshold, then the individual is accepted and treated as a nestmate, whereas non-nestmate rejection occurs if above this threshold (Lacy and Sherman 1983; Reeve 1989). Recognition labels are predicted to be multi-component, variable phenotypes (Leonhardt et al. 2016; Tibbetts et al. 2020). Therefore, recognition labels can be imagined as signature mixtures, rather than one specific signal (Dani et al. 2001; Wyatt 2010; Elia et al. 2017). More generally, the best signals are predicted to be those that reliably and robustly activate the nervous system of the receiver (Endler 1992). The importance of odor in social insect recognition processes was appreciated by perspicacious researchers like Adele Fielde prior to the major development of instruments for analytical chemistry (Fielde 1905). Understanding the chemical composition and development of the recognition label used during nestmate

recognition is a priority for research on social insects (Page et al. 1991; Espelie et al. 1995; Dani et al. 2001; Vernier et al. 2019).

Cuticular hydrocarbons (CHCs) are chemicals used by insects to waterproof the cuticle and to convey information to conspecifics (Blomquist & Bagnères 2010). CHCs are typically large compounds with high molecular weight and low polarity, making them a relatively stable phenotype and a reliable source of information over time (Martin et al. 2009). Nestmate recognition in diverse social insect species depends in-part on CHCs as recognition signatures (Hölldobler and Michener 1980; Bonavita-Cougourdan et al. 1987; Singer 1998). As with other traits used by animals to recognize each other, chemical nestmate discrimination requires high intraspecific variation in recognition phenotypes, and CHC profiles must be sufficiently variable within a site for nests to be distinguishable (van Zweden JS and d'Ettorre 2010). Predictions for the role of CHCs in nestmate recognition emerged from correlational studies identifying nest-specific patterns of variation in CHCs (Espelie et al. 1994; Layton et al. 1994; Martin et al. 2008a). Behavioral assays supported this hypothesis (Singer and Espelie 1996; Lorenzi et al. 1997), and manipulative experiments confirmed the function of CHCs as nestmate recognition signature mixtures in ants, wasps, and bees (Breed and Stiller 1992; Lahav 1999; Dani et al. 2001; Martin et al. 2008b).

Methyl-branched CHCs (methylalkanes) and alkenes have been hypothesized to be especially suitable as signal phenotypes because of their characteristic stereochemistry, which likely facilitates their detection by the insect olfactory system (Espelie et al. 1994; Van Zweden and d'Ettorre 2010; Pask et al. 2017; but see di

Mauro et al. 2015). For example, small 2-carbon differences in the position of the methyl branch in methylnonacosane (MeC29) resulted in distinct responses from heterologously expressed odorant receptors from the ant *Harpegnathos saltator* (Pask et al. 2017). The variable stereochemistry of these compounds is detectable by insects, as insect olfactory systems can discern enantiomers (Niehuis et al. 2013; Sims et al. 2022). Behavioral assays in *Polistes dominula* paper wasps supported the role of two methylalkanes and two methylalkenes but not six linear alkanes in nestmate recognition (Dani et al. 2001). Argentine ants *Linepithema humile* more readily learned to recognize tri-methylalkanes than mono-methylalkanes or linear alkanes in conditioning experiments (Van Wilgenburg et al. 2012).

Paper wasps in the *Polistes* genus cooperate with nestmates and aggressively deter non-nestmates when they are encountered at the nest (Gamboa et al. 1986a; Legan 2022). In *Polistes*, the nestmate recognition signature consists of cuticular hydrocarbon chemicals (CHCs) on the cuticle and the nest (Gamboa et al. 1986a; Starks et al. 1998; Singer and Espelie 1992; Dani et al. 2001). A common feature of *Polistes* CHC profiles is the frequency of methylalkanes (Dani et al. 2006; Loope 2015). In *P. exclamans*, the most informative features distinguishing workers from two different nests were the proportions of methylheptacosanes (MeC27) and methylnonacosanes (MeC29) (Singer et al. 1992). Treatment of *P. dominulus* workers with methylalkanes and methylalkenes elicited more aggression from nestmates than treatment with linear alkanes (Dani et al. 2001). Observations of multiple paper wasp species provide evidence for the nestmate recognition signature function of methylalkanes (Bonavita-Cougourdan et al. 1991; Espelie et al. 1994).

In the northern paper wasp *P. fuscatus*, nestmate recognition has been observed in spring foundresses, workers, gynes, and males, suggesting the ability to discriminate between nestmates and non-nestmates is important in diverse behavioral contexts (foundresses: Post & Jeanne 1982; Bornais et al. 1983; workers: Pfennig et al. 1983a; gynes: Gamboa et al. 1985; Shellman & Gamboa 1982; Pfennig et al. 1983b; males: Ryan & Gamboa 1986, but see Ryan et al. 1984; de Souza et al. 2017; de Souza et al. 2021). While also capable of individual facial recognition, *P. fuscatus* rely on CHCs for nestmate recognition (Ortiz and Tibbetts 2020). Methylalkanes are especially variable between nests of *P. fuscatus*, supporting their role as recognition compounds (Espelie et al. 1994). Espelie et al. (1994) proposed criteria for nestmate recognition signature compounds and identified 13- and 15-MeC31, 11,15- and 13,17-diMeC31, and 13-, 15-, and 17-MeC33 as likely recognition compounds. In a follow-up study combining chemical analysis with behavioral assays of *P. fuscatus* paper wasps, differences in the abundance of three methylalkanes (including two of the methylalkanes highlighted by Espelie et al. 1994), an alkene, and six linear alkanes predicted rejection of a wasp from the nest (Gamboa et al. 1996). These studies emphasized that other CHCs undetected by their methods may play a role in the nestmate recognition signature of *P. fuscatus*. More chemical ecology analyses using newer instruments are needed to identify candidate recognition signature CHCs of *P. fuscatus*, which can then be used in behavioral assays to determine their impact on recognition.

By studying *P. fuscatus* colonies in the field during two consecutive summers at the Liddell Field Station in Ithaca, NY, hydrocarbon data from 189 wasps from 26 nests were collected. We identified CHCs, investigated nest-specificity of these compounds,

and explored how compounds are co-expressed. Our results expand the list of putative nestmate recognition signature compounds in *P. fuscatus*.

Methods

Study organisms and study site:

Forty *Polistes fuscatus* nests and associated foundresses were collected from one site in Ithaca, NY (42°31'59.1" N, 76°18'47.0" W) before 8AM EST on May 24, 2019. Nests were temporarily housed in deli cups and brought to the lab where they were provided sugar cubes and cotton-plugged glass vials filled with water *ad libitum*. After a short period of cold anesthetization in a refrigerator (~4-6°C), wasps were removed from their nests and given individual specific paint marks or QR codes glued to the mesoscutum (dorsal thorax). The nest pedicel was hot glued to the ceiling of a modified bird box. Wasps were returned to the nest, and the metal mesh face of the box was covered with a fine plastic mesh to prevent the wasps exiting the box. The plastic mesh was removed after three to five days of acclimation to the nest box. Boxes were fixed to wooden posts at the Liddell Field Station in Ithaca, NY (42°27'36.7" N, 76°26'39.2" W). About half of the nests continued to develop after transplant, and by June 18th, 2019, there were 23 active *P. fuscatus* nests under observation in the field at Liddell Field Station, including three nests that were naturally established at the site. Nests maintained in the lab were on a 16L:8D light cycle. In the summer and fall 2019, 182 female offspring of the 2019 foundresses were marked with colony-specific paint marks. Of these, 22 were gynes that returned to found nests of their own in the spring of 2020.

Chemical sampling using SPME and hexane washes:

Contact solid-phase microextraction was performed using 100 μm PDMS SPME fibers, which have been used previously to sample cuticular hydrocarbons by contact with insect cuticle (Ferreira-Caliman et al. 2012; Ferreira-Caliman et al. 2014; Gerhardt et al. 2015; Bonacci et al. 2018). SPME fibers (MilliporeSigma SPME Portable Field Sampler, SKU: 504823) sampled the nonpolar and high MW compounds with molecular weight between 125-600 amu on the wasp cuticle (Liu & Ouyang 2017). Twelve fibers were individually numbered and conditioned according to the manufacturer's instructions. Wasps were sampled during a quiescent activity period, between 10PM and 4AM EST. Nest boxes were removed from their posts in the field and covered with a black mesh bag before being placed in a refrigerator (~4-6°C). Cold anesthetized wasps were removed from the nest by hand, using nitrile gloves, and CHCs were sampled by swabbing the dorsal abdominal cuticle (tergites 2 and 3) 40 times with a SPME fiber (video S1). The abdomen was sampled because CHCs are abundant on this body part in Hymenoptera (Monnin et al. 1998). If a wasp lacked an identifying label (QR code or paint mark), then it was painted on the mesoscutum and sometimes the wings. Each wasp was handled using a fresh pair of nitrile gloves. After a second bout of cold anesthetization, the nest box was returned to its post in the field at the Liddell Field Station. SPME fibers were stored in SUPELCO field sampler holders, tip down, with tips of field sampler holders resting in 15 mL polypropylene tubes in a tube rack until injection into the GC-FID. Compound identities were determined for three hexane-washed female *P. fuscatus* collected from a site in Ithaca, NY

(42°25’49.67” N, 76°30’11.13” W) and analyzed using gas chromatography-mass spectrometry as described elsewhere (Loope 2015).

Chemical detection and peak integration:

The chemicals adsorbed by SPME fibers were analyzed using a Shimadzu 2014 gas chromatography-flame ionization detection (GC-FID). A DB-5 capillary column (length: 30 m; diameter: 0.25 mm; film: 0.25 μm) was used after it was conditioned according to the manufacturer’s instructions. The SPME fiber was manually injected by fully extending the fiber into the injection port, where chemical samples were desorbed for five minutes at 270°C or 280°C. The inlet pressure was 96.4 kPa and column flow rate was 1.2 ml/min for a linear velocity of 29.1 cm/s. The oven temperature was held at 50°C for two minutes and ramped to 220°C at a rate of 30°C per minute. To increase separation between peaks in the range of C20-C40, the rate was decreased to 5°C per minute until the oven temperature reached 290°C, then further decreased to 3°C per minute until the oven temperature reached 320°C, where it was held for 12.5 minutes. The total program time per sample was 44.17 min. Peaks were integrated in Shimadzu LabSolutions version 5.82 with a minimum slope threshold of 1000 μV/s. Raw data was exported from LabSolutions as ASCII files and integrated peaks were exported from LabSolutions as peak tables in Microsoft Excel files. Protocol for GC-MS to determine compound identities in representative *P. fuscatus* samples is described in Loope (2015).

Peak alignment and compound inference:

Peak tables were imported from separate sheets of a Microsoft Excel file as a list of data frames to R (R Core Team 2018). Retention indices (RI) of known linear

alkane standards were added to standard peak tables, where RI of linear alkane standards is equal to the number of carbons in the linear alkane multiplied by 100. Linear relationships between retention times and retention indices were confirmed by fitting a linear model in R using the function “lm”. Using the linear alkane standards analyzed on the nearest calendar date to each sample run, retention times of standards were regressed against retention indices of standards, and retention indices corresponding to retention times of unknown peaks in samples were interpolated using the “approx” function and added as a column to the peak table. The list of indexed peak tables was used as input for R package GCalignR (Ottensmann et al. 2018). GCalignR was run with options “rt_col_name = ‘RI’” “rt_cutoff_low = 2080” “rt_cutoff_high = 3780” “max_linear_shift = 7” “max_diff_peak2mean = 6” “min_diff_peak2peak = 7”. To filter out artifactual qualitative variation across samples resulting from alignment errors, the two GCalignR alignments were inspected (2019: 174 peak calls; 2020: 155 peak calls) and peaks were merged in cases where the two peaks were (1) within ≤6 RI units of each other, and (2) were completely complementary in all but ≤6 samples. When merging, the peak area of the less common peak was added to the peak area of the more common peak, and the less common peak was removed. Merging resulted in the removal of 26 peaks in the 2019 dataset and 15 peaks in the 2020 dataset. For each year’s alignment, only those peaks present in ≥50% of samples were retained for analyses, ultimately retaining 72 peaks from 2019 and 70 peaks from 2020.

To determine putative chemical compositions of peaks, retention indices of peak tables generated using GC-MS were manually aligned to the mean retention indices from the unfiltered 2019 and 2020 alignments. When a peak from the SPME field

dataset could not be identified based on the GC-MS peaks, it was retained and included in the list as unknown ("UNK") if present in at least half of individuals sampled during at least one field season. This process identified 5 instances of split peaks to be merged into one peak (2019: 5-MeC25, 4-MeC28, diMeC30; 2020: C21, diMeC35). Only peaks with putative identities determined by GC-MS and present in at least half of samples per field season were retained for clustering analyses (2019: 58 peaks; 2020: 57 peaks).

Statistical analyses:

Chromatograms from 95 post emergence (worker or gyne) females across 13 nests sampled in 2019, and 94 post emergence females across 13 nests sampled in 2020 were analyzed. All analyses were performed separately for each year's dataset. All analyses were performed in R version 4.1.2 (R Core Team 2018). Non-metric multidimensional scaling (NMDS) analyses were conducted using the "metaMDS" function with Bray-Curtis dissimilarity index and "trymax = 50" in the R package vegan version 2.6-2 (Oksanen et al. 2022). NMDS stress values were tolerable (2019: 0.114; 2020: 0.142). Analyses of similarities (ANOSIM) were performed using the vegan function "anosim" with 9999 permutations and Bray-Curtis dissimilarity index. Discriminant analyses of principal components (DAPC) were conducted using the R package adegenet version 2.1.5 (Jombart 2008; Jombart et al. 2010).

DAPC was conducted on the log(x+1) transformed relative abundance matrices (Brückner and Heethoff 2017). The cross-validation function "xvalDapc" was used to select the number of principal components to retain during preliminary DAPC for all analyses. The preliminary DAPC was run and then used as input to the "optim.a.score"

function to determine the optimal number of PCs to retain to balance overfitting with error, sometimes indicating an optimal number of PCs lower than that originally identified using "xvalDapc." If a lower optimal number of PCs was indicated, then the DAPC was rerun with the lower PCs setting to avoid overfitting the discriminant functions (Jombart and Collins 2015). Number of discriminant functions retained was the maximum possible for each analysis (number of nests (13) – 1 = 12). For the 2019 data, DAPC of individuals from 16 nests revealed extreme outliers from one nest from a different source population (19S, founded by a local wasp) and from a nest reared in the lab (19Lab5) (fig 1.1).

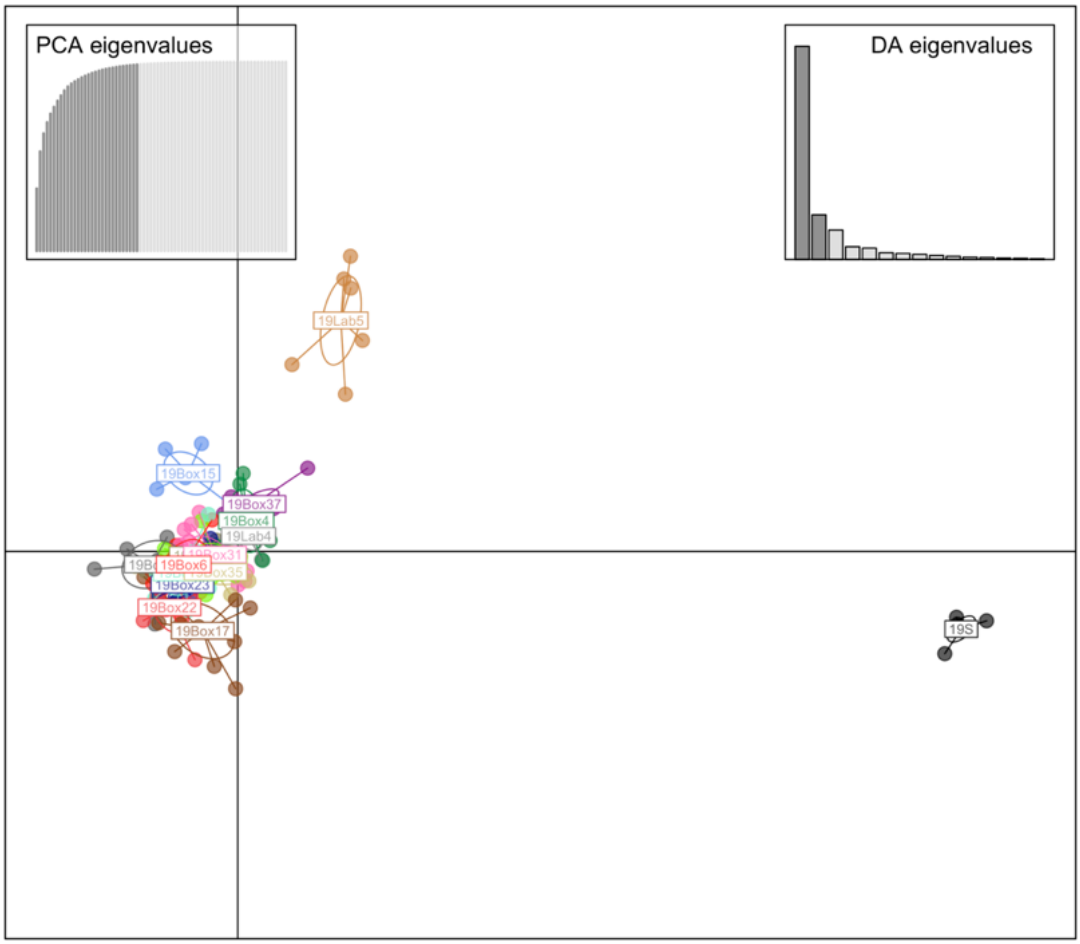

Figure 1.1. Scatter plot of first two discriminant functions, results of DAPC on individuals sampled in 2019. Outlier nests are "19S" (black) and "19Lab5" (orange).

Therefore, individuals from these nests and another nest reared in the lab (19Lab4) were removed from the dataset for subsequent analyses, since our focus concerned within site variation in CHCs. DAPC (18 PCs retained) on the remaining 95 wasps from 13 nests did not reveal extreme outliers (fig. 1.3C). For the 2020 data, DAPC (13 PCs retained) of 94 individuals from 13 nests did not indicate extreme outliers (fig. 1.3D). Function "loadingplot" was used to view the top peaks explaining the variation in the first two discriminant functions. Wasps were assigned to 13 random groups and DAPC performed with random group as factor to generate as a negative control. Random numbers were generated using the "sample(1:13, X, replace=T)" function in R, where "X" was 95 and 94 for 2019 and 2020, respectively.

Random Forest classification was carried out using R package randomForest version 4.7.1.1 with "mtry=8", "importance=T", "ntree=10000" and variable importance was assessed using function "importance" (Liaw and Wiener 2002).

Spearman's rho values were calculated with function "rcorr" from R package Hmisc version 4.7.0 (Harrell 2019). Heatmaps were generated using R package ggplot2 version 3.3.6 (Wickham 2016).

Results

Chemical datasets:

96 peaks were detected with SPME sampling and GC-FID in the summers of 2019 and 2020 (fig. 1.2). 81 were identifiable based on GC-MS data and 15 were unidentifiable (table S1.1). The 20 peaks detected by Espelie et al. (1994) in *P. fuscatus* were recovered in our chromatography. CHCs documented by Espelie et al. (1994) were typically highly abundant relative to other peaks within chromatograms.

27 peaks contained mono-methylalkanes, 22 contained di-methylalkanes, 14 contained tri-methylalkanes, 11 contained linear alkanes, 4 contained tetra-methylalkanes, 3 contained methyl alkenes, and 3 contained linear alkenes.

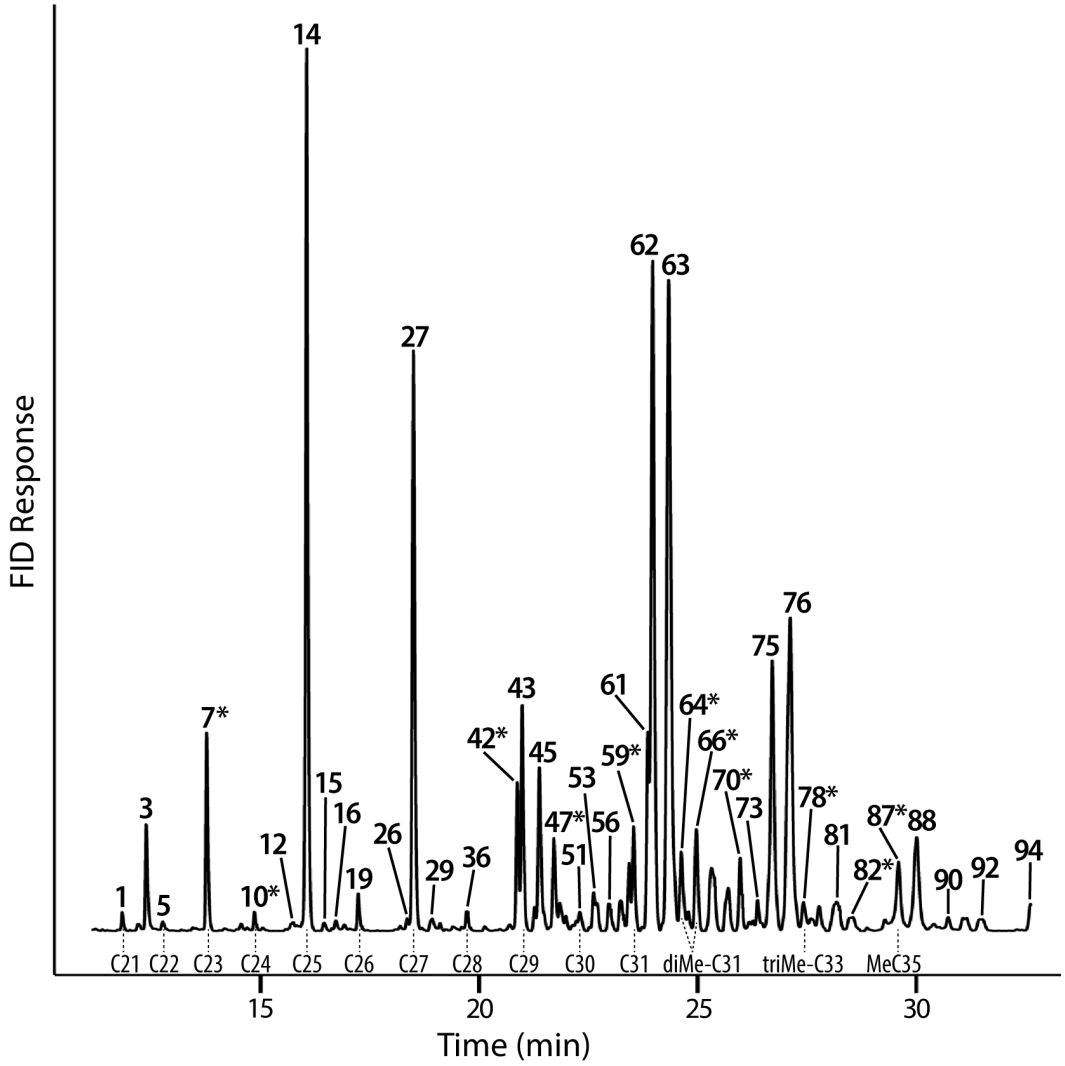

Figure 1.2. Representative chromatogram of the solid-phase microextract from a female *P. fuscatus* sampled in 2019. Select peaks are numbered corresponding to their numbering in table S1.1. 11 peaks especially diagnostic of nest membership, based on evidence from this study and previously published work, are indicated by an asterisk and listed in table 1.1. All linear alkane (C21 – C31) as well as three methylalkane peaks are indicated at the base of the chromatogram as a reference to infer the chain length of nearby peaks.

Cuticular hydrocarbon mixtures are diagnostic of nest affiliation:

Wasps clustered by nest with some overlap in NMDS analyses (fig. 1.3A, 1.3B). Analyses of similarities (ANOSIM) rejected the null hypothesis that there is no difference between CHC profiles between nests, but R values near 0.3 indicate some overlap in CHC profiles between nests (2019: R = 0.2561, P = 1e-04; 2020: R = 0.2951, P = 1e-04). To identify the peaks driving nest specific CHC profile variation, DAPC with nests as a priori groups was used, because DAPC reveals variables that maximize the separation between groups (fig. 1.3C, 1.3D). DAPC assigned individuals sampled in 2019 to their correct nest with 78% accuracy, a 31% increase in accuracy compared to random chance. DAPC assigned individuals sampled in 2020 to their correct nest with 80% accuracy, a 41% increase in accuracy compared to random chance.

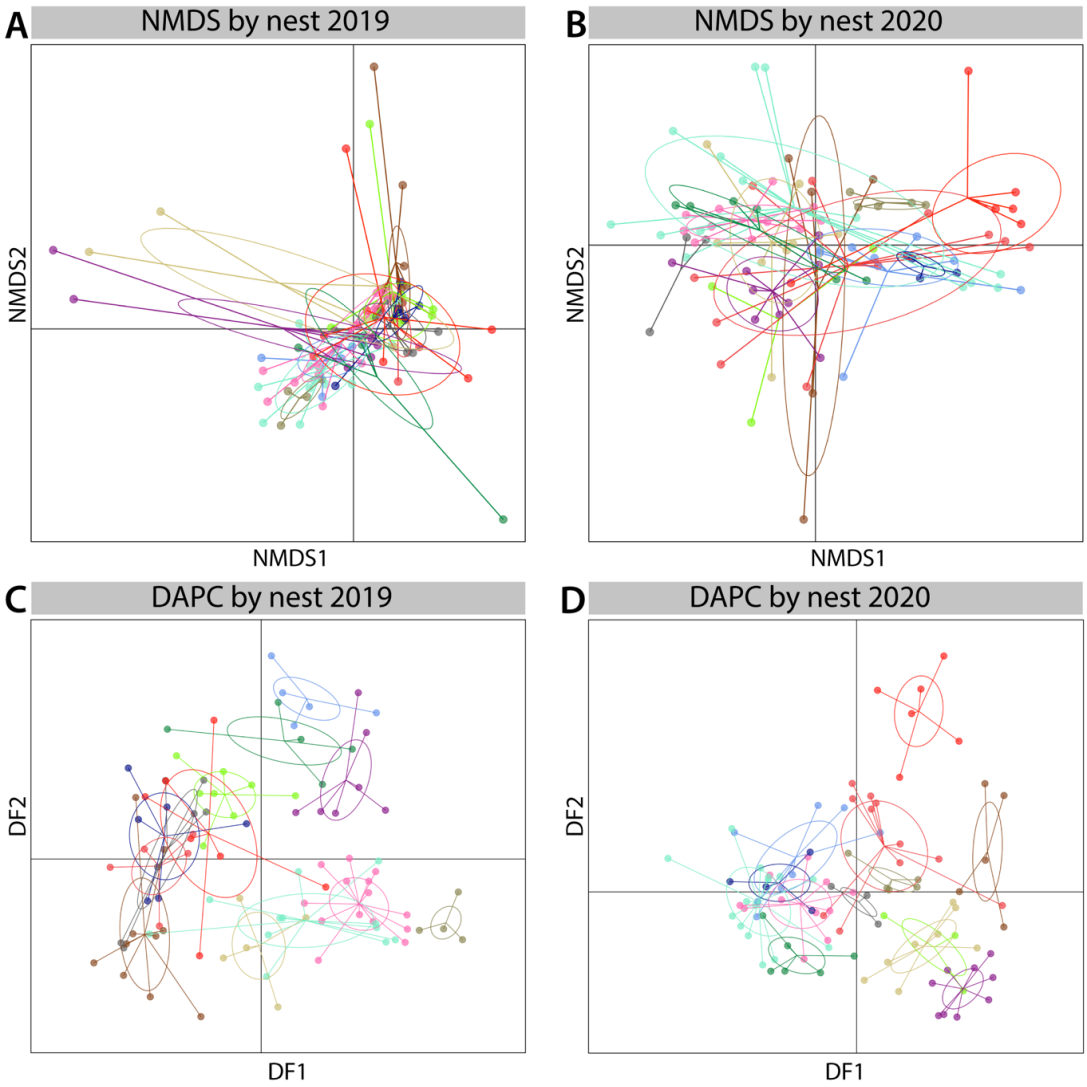

Figure 1.3. Results of multivariate ordination and clustering analyses of 95 CHC profiles from 2019 and 94 from 2020. Different nests (13 in each year) are coded by different colors. (A) Plot showing individual CHC profiles represented as coordinate pairs along two non-metric multidimensional scaling (NMDS) axes using Bray-Curtis dissimilarities. (B) Plotting 2020 data on two NMDS axes. (C) Plot showing individual CHC profiles represented as coordinate pairs of two discriminant functions generated using Discriminant Analysis of Principal Components (DAPC) for 2019 data. (D) Plotting 2020 data on two discriminant functions using DAPC.

Top canonical loadings from DAPC of 2019 data correlated 9,13,17-triMeC33 with DF1, and 5,11,19-triMeC35 with DF2 (fig. 1.4A-B). From DAPC of 2020 data, top canonical loadings correlated C31 with DF1 and 9-, 11-, 13-, and 15-MeC31 with DF2 (fig. 1.4C-D).

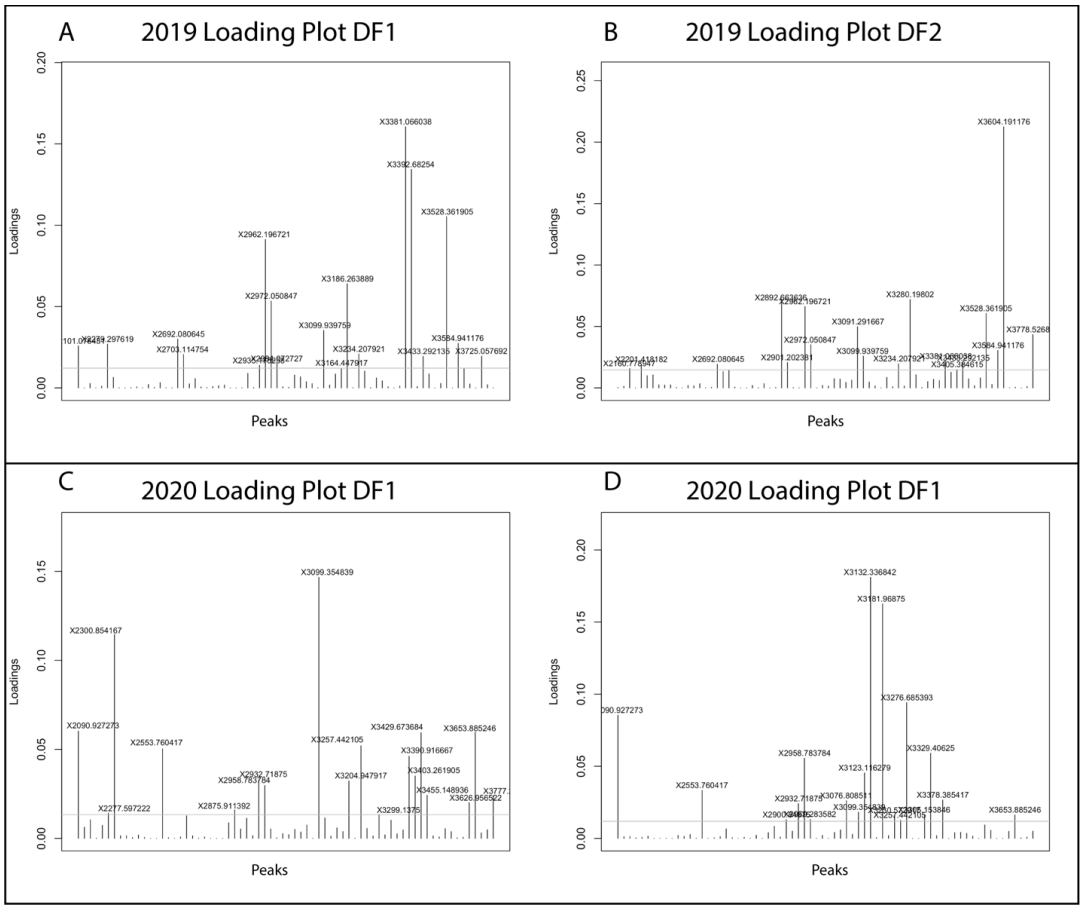

Figure 1.4. Canonical loadings correlating peaks (as retention indices) with first two discriminant functions in DAPC analyses of 2019 (A, B) and 2020 (C, D) data.

For 72 peaks from the 2019 dataset, linear models were fitted for each peak with peak relative abundance as the dependent variable and nest identity as the independent variable, and ANOVA was run on the model. 24 peaks were significantly different between nests after Bonferroni correction ($P < 0.00035$). ANOVA of 70 peaks from 2020 samples indicated 28 peaks significantly different between nests. Finally, random forest (RF) analysis was used to determine the top three peaks contributing to accurate nest assignment in each year. Nest-specificity of eleven compounds was supported by evidence from at least three of the following sources: DAPC, RF, ANOVA, and either Espelie et al. (1994) or Gamboa et al. (1996) (table 1.1).

Table 1.1. **List of peaks especially diagnostic of nest affiliation:** Eleven peaks diagnostic of nest affiliation present in ≥90% of individuals across years. These peaks are designated by asterisks in fig. 1.2.

		2019 evidence			2020 evidence		
#	Compound(s)	DAPC	RF	ANOVA	DAPC	RF	ANOVA
7	n-C23[!]				*		*
10	n-C24[!]			*			*
42	C29:1[!]	*					*
47	11,15-diMeC29[!]			*			*
59	n-C31[!]			*	*		*
64	5,13-diMeC31 5,15-diMeC31 5,17-diMeC31 5,19-diMeC31 5,21-diMeC31 5,23-diMeC31		*		*		*
66	3,15-diMeC31 3,17-diMeC31 3,19-diMeC31 3,21-diMeC31 5,9,21-triMeC31			*		*	*
70	UNK_07	*		*	*		
78	9,13,17-triMeC33	*	*	*			*
82	3,7,11,19-tetraMeC33			*		*	*
87	11-MeC35 13-MeC35 15-MeC35 17-MeC35	*	*	*		*	*

! indicates the compound was identified by Espelie et al. (1994) as diagnostic of nest affiliation or Gamboa et al. (1996) as associated with aggression directed towards non-nestmates.

While methylalkanes represented the top nest-specific peaks in random forest analysis, the mean nest-specificity, as measured by mean decrease in accuracy, between compound classes was not significantly different (ANOVA: 2019, $F = 1.275$, $P = 0.2888$; 2020, $F = 0.212$, $P = 0.956$) (fig. 1.5B, 3D; fig. 1.6).

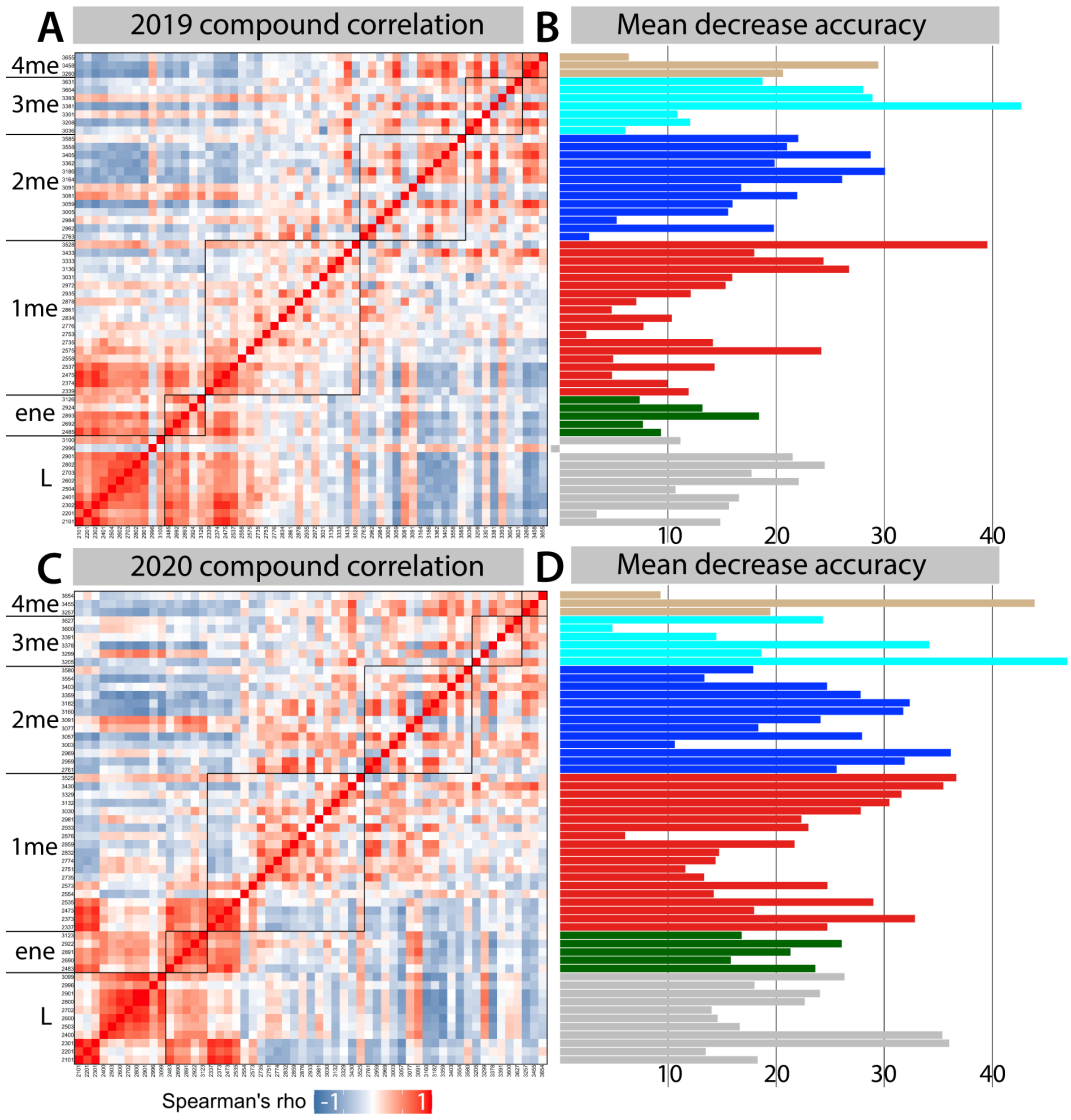

Figure 1.5. (A) Heat map showing correlations between peak relative abundances in wasps sampled in 2019. Peaks are ordered by chemical class (4me: tetra-methylalkanes; 3me: tri-methylalkanes; 2me: di-methylalkanes; 1me: mono-methylalkanes; ene: unsaturated compounds; L: linear alkanes) and within classes peaks are ordered by retention index (increasing along axes). Boxes bound correlations between peaks in the same chemical class. (B) For each peak in (A), the mean decrease in model accuracy in assigning individuals to correct nest when the peak is removed during random forest analysis. (C) Heat map showing correlations between peak relative abundances in wasps sampled in 2020. Boxes bound correlations between peaks in the same chemical class. (D) For each peak in (C), the mean decrease in model accuracy in assigning individuals to correct nest when the peak is removed during random forest analysis.

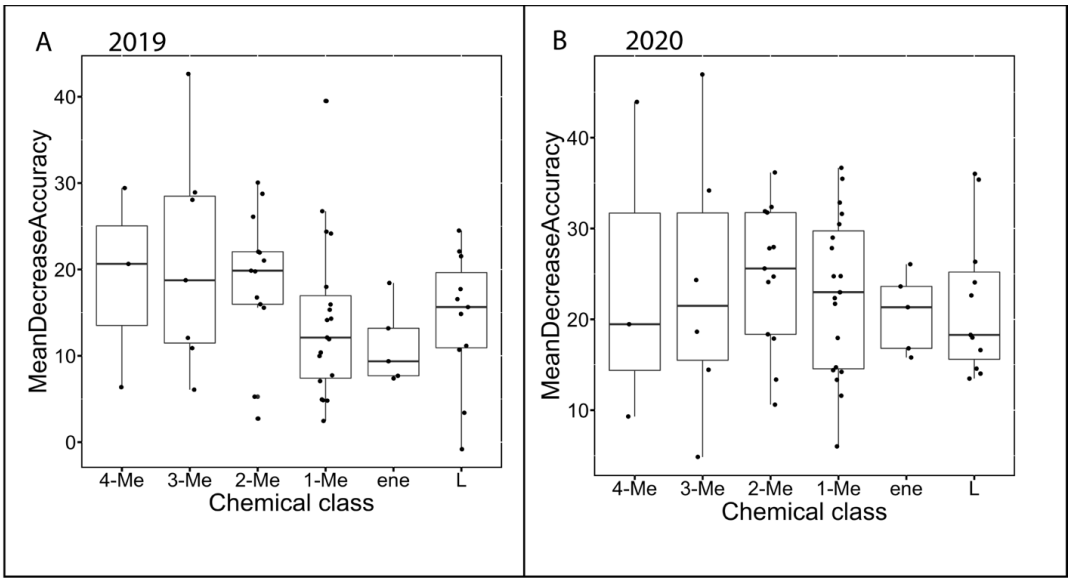

Figure 1.6. Box and whisker plots displaying mean decrease in accuracy per compound in random forest analyses by chemical class for (A) 2019 and (B) 2020.

Correlations and anticorrelations in compound abundance:

Strong correlations between peaks would suggest they carry similar information and are not independent units in the nestmate recognition signature mixture. Considering the peaks listed in table 1.1, strong positive correlations ($\rho > 0.5$) were found between one pair of peaks. between diMeC31 (peak 64) correlated with 11,15-diMeC29 (peak 47) ($\rho = 0.80$). Therefore, these two di-methylalkanes may carry redundant information about nest affiliation.

The top four positive correlations between 2019 peaks were between C21 (peak 1) and C23 (peak 7*) ($\rho = 0.95$), C25:1 & 2-MeC24 (co-eluting in peak 12) and C21 (peak 1) ($\rho = 0.90$), triMeC33 & diMeC33 (co-eluting in peak 80) and diMeC31 & triMeC31 (co-eluting in peak 66*) ($\rho = 0.90$), tetraMeC31 & diMeC32 (co-eluting in peak 68) and diMeC30 (peak 56) ($\rho = 0.89$). The top negative correlations between 2019 peaks were between multi-methyl branched alkanes and linear alkanes: specifically, between diMeC30 (peak 56) and C23 (peak 7*) ($\rho = -0.77$), between

tetraMeC31 & diMeC32 (co-eluting in peak 68) and C23 (peak 7*) (ρ = -0.76), between 9,13,17-triMeC33 (peak 78*) and C29 (peak 43) (ρ = -0.72), and between diMeC31 (peak 63) and C29 (peak 43) (ρ = -0.72).

The top four positive correlations between 2020 peaks were between C28 (peak 36) and C29 (peak 43) (ρ = 0.97), C21 (peak 1) and C23 (peak 7*) (ρ = 0.95), C21 (peak 1) and MeC23 (peak 8) (ρ = 0.91), and between C26 (peak 19) and C28 (peak 36) (ρ = 0.91). The top negative correlations between 2020 peaks were between C28 (peak 36) and diMeC33 (peak 76) (ρ = -0.79), C26 (peak 19) and diMeC33 (peak 76) (ρ = -0.79), C29 (peak 43) and diMeC31 (peak 63) (ρ = -0.77), and between C31 (peak 59*) and diMeC31 (peak 63) (ρ = -0.77).

Discussion

The results presented here support the hypothesis that CHC profiles are signature mixtures in *Polistes* paper wasps. Compared to pheromones, signature mixtures are learned (Wyatt 2010). Since CHC profiles function as signature mixtures in social insects, there may not be strong stabilizing selection favoring expression of any specific component of the mixture at a given time during the colony cycle. Therefore, the composition of the nest signature mixture is free to fluctuate over time if the mixture remains variable and detectable. This is expected over evolutionary time, as well as from season to season, and within the course of a season. This may explain why, while some compounds were highly diagnostic of nest in multiple sites and seasons, others appeared to vary in nest-specificity between years and between studies (Espelie et al. 1994; Gamboa et al. 1996). Future work should examine how the *P. fuscatus* CHC profile varies seasonally.

In this study, multiple methylalkanes were especially diagnostic of *P. fuscatus* paper wasp nest membership, replicating previous findings of nest-specific methylalkane ratios in *P. fuscatus* and other Polistine wasps (Espelie et al. 1994). In *P. dominula*, while treatment of nestmates with linear alkanes had no effect on tolerance, treatment with methylalkanes and methylalkenes reduced tolerance from nestmates (Dani et al. 2001). In the Polistine wasp *Mischocyttarus cassununga*, concentrations of mono-, di-, and tetra-methylalkanes were significantly different across nests (Murakami et al. 2015). Across ants, di-methylalkanes are highly species- and colony-specific and represent the most diverse class of CHCs (Dahbi et al. 1996; Martin & Drijfhout 2009). For example, in the ant *Formica fusca*, nine positional isomers of diMeC25 encode colony affiliation (Martin et al. 2008). Compounds listed in table 1.1 would be predicted to be especially attended to during nestmate recognition by wild *P. fuscatus* wasps.

In addition to contributing to the nest signature mixture, methylalkanes also function as signals and cues of fertility in social insects (Sledge et al. 2001; Sledge et al. 2004; Liebig 2010; Van Oystaeyen et al. 2014; Kingwell 2020; Kingwell et al. 2021). In *Polistes satan*, compounds strongly associated with fertility included linear alkanes and methyl alkanes (Oi et al. 2019). In at least one *Polistes* species, *P. gallicus*, hydrocarbons associated with queens and those associated with nest membership were secreted by the ventral abdominal Van der Vecht organ (Dapporto et al. 2007). Since *Polistes* paper wasps are sensitive to changes in the relative abundance of CHCs when discriminating nestmates (Dani et al. 2001; Van Zweden JS and d'Ettorre 2010) and

given that females within a nest vary in fertility status, fertility correlated CHCs would not be expected to perform a dual function in the nest recognition signature.

As in the case of human fingerprints, phenotypic variation does not necessarily imply the presence of meaningful phenotypic variation from the standpoint of the receiver's sensory system. The CHC fraction is determinedly the nest signature recognition mixture in *P. fuscatus* (Ortiz and Tibbetts 2020). An important next step will be to assess the responses of *P. fuscatus* wasps to experimentally altered CHC profiles following the methods of Dani et al. (2001). *Polistes* CHC profiles may be influenced by the environment, and wasps can acquire cuticular chemicals from the nest paper (Gamboa et al. 1986b; Pickett et al. 2000). However, the prevalence of environmental hydrocarbon acquisition by wild paper wasps has yet to be determined. A significant heritable component of nest CHC profiles was recently found in the ant *Monomorium pharaonis* (Walsh et al. 2020). Future work should quantify CHC profile heritability in *P. fuscatus*.

Espelie et al. (1994) highlighted three *P. fuscatus* CHCs that best matched criteria for recognition signature compounds: 13- & 15-MeC31, 11,15- & 13,17-diMeC31, and 13-,15-, & 17-MeC33, and we found evidence from univariate or multivariate analyses supporting this claim for each of these compounds. Gamboa et al. (1996) identified ten peaks predictive of nestmate recognition behavior, and our analyses found evidence supporting the possible function of all of these in the signature mixture. By detecting more of the *P. fuscatus* CHC profile, this study uncovered 28 additional peaks exhibiting nest-specificity, providing a comprehensive list as well as a

more targeted list (table 1.1) of compounds that should be queried in nestmate recognition behavioral assays in the future.

Acknowledgments

This work was supported by the North American Section of the International Union for the Study of Social Insects Robert L. and Louise B. Jeanne Social Wasp Research Grant (to AWL), National Science Foundation Graduate Research Fellowship Program (Grant No. DGE-1650441 to AWL), National Science Foundation CAREER Grant (grant number DEB-1750394 to MJS), and National Institutes of Health grant number DP2-GM128202 (to MJS). GC-MS data for three *Polistes fuscatus* wasps was generously provided by Dr. Kevin Loope and Dr. Patrizia d'Ettorre (Loope 2015). AWL thanks Ethan Bass for sharing useful R scripts for linear interpolation of retention times to generate retention indices, and Callum Kingwell for help troubleshooting SPME sampling methods. The Cornell Liddell Field Station made this work possible.

REFERENCES

Arnold G, Quenet B, Cornuet J-M, Masson C, DeSchepper B, Estoup A, Gasquil P. 1996. Kin recognition in honeybees. Nature. 379:498.

Arthur CL, Pawliszyn J. 1990. Solid phase microextraction with thermal desorption using fused silica optical fibers. Analytical Chemistry 62(19):2145-2148.

Blomquist, G.J., Bagnéres,A.-G. 2010. *Insect Hydrocarbons: Biology, Biochemistry, and Chemical Ecology*. Cambridge University Press: New York, USA.

Bonacci, T., Mazzei, A., Naccarato, A., Elliani, R., Tagarelli, A., Brandmayr, P. 2018. Beetles "in red": are the endangered flat bark beetles *Cucujus cinnaberinus* and *C. haematodes* chemically protected? (Coleoptera: Cucujidae). *The European Zoological Journal* 85(1):128-136.

Bonavita-Cougourdan A, Clément J-L, Lange C. 1987. Nestmate recognition: the role of cuticular hydrocarbons in the ant *Camponotus vagus* Scop. Journal of Entomological Sciences. 22:1-10.

Bonavita-Cougourdan A, Theraulaz G, Bagnères M, Roux M, Pratte M, Provost E, Clément J-L. 1991. Cuticular hydrocarbons, social organization and ovarian development in a Polistine wasp: *Polistes dominulus* Christ. Computational Biochemical Physiology B. 100:667-680.

Bornais KM, Larch CM, Gamboa GJ, Daily RB. 1983. Nestmate discrimination among laboratory overwintered foundresses of the paper wasp, *Polistes fuscatus* (Hymenoptera: Vespidae). Canadian Entomologist 115:655-658.

Breed MD, Stiller TM. 1992. Honey bee, *Apis mellifera*, nestmate discrimination: hydrocarbon effects and the evolutionary implications of comb choice. Animal Behaviour 43(6):875-883.

Breed MD, Deng X-B, Buchwald R. 2007. Comparative nestmate recognition in Asian honey bees, *Apis florea*, *Apid andreniformis*, *Apis dorsata*, and *Apis cerana*. Apidologie. 38:411-418.

Brückner A, Heethoff M. 2017. A chemo-ecologists' practical guide to compositional data analysis. Chemoecology 27:33-46.

Carlson DA, Bernier UR, Sutton BD. 1998. Elution patterns from capillary GC for methyl-branched alkanes. Journal of Chemical Ecology 24(11):1845-1865.

Chung H, and Carroll SB. 2015. Wax, sex and the origin of species: dual roles of insect cuticular hydrocarbons in adaptation and mating. BioEssays 37:822-830.

Dahbi, A., Lenoir, A., Tinaut, A., Taghizadeh, T., Francke, W., Hefetz, A. 1996. Chemistry of the postpharyngeal gland secretion and its implication for the phylogeny of Iberian *Cataglyphis* species (Hymenoptera: Formicidae). Chemoecology 7:163-171.

Dani FR, Jones GR, Destri S, Spencer SH, Turillazzi S. 2001. Deciphering the recognition signature within the cuticular chemical profile of paper wasps. Animal Behaviour. 62(1):165-171.

Dani FR. 2006. Cuticular lipids as semiochemicals in paper wasps and other social insects. Annales Zoologici Fennici 43:500-514.

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.

De Souza AR, Barbosa BC, da Silva RC, Prezoto F, Lino-Neto J, do Nascimento FS. 2017. No evidence of intersexual kin recognition by males of the neotropical paper wasp *Polistes versicolor*. Journal of Insect Behavior. 30:180-187.

De Souza AR, da Silva RC, Batista NR, Júnior WFA, do Nascimento FS. 2021. Lack of caste discrimination by males during sexual context in a neotropical paper wasp. Ethology. 127(7):613-619.

Dembeck, L.M., Böröczky, K., Huang, W., Shal, C., Anholt, R.R.H., Mackay, T.F.C. 2015. Genetic architecture of natural variation in cuticular hydrocarbon composition in *Drosophila melanogaster*. eLife 4:e09861.

Di Mauro G, Perez M, Lorenzi MC, Guerrieri FJ, Millar JG, d'Ettorre P. 2015. Ants discriminate between different hydrocarbon concentrations. Frontiers in Ecology and Evolution 3(133):1-11.

Elia M, Blancato G, Picchi L, Lucas C, Bagnères A-G, Lorenzi MC. 2017. Nest signature changes throughout colony cycle and after social parasite invasion in social wasps. PLOS ONE 12(12):e0190018.

Endler JA. 1992. Signals, signal conditions, and the direction of evolution. The American Naturalist. 139:S125-S153.

Espelie KE, Gamboa GJ, Grudzien TA, and Bura EA. 1994. Cuticular hydrocarbons of the paper wasp *Polistes fuscatus*: a search for recognition pheromones. Journal of Chemical Ecology 20(7):1677-1687.

Ferreira-Caliman MJ, Zucchi R, Nascimento FS. 2012. Cuticular hydrocarbons discriminate distinct colonies of *Melipona marginata* (Hymenoptera, Apinae, Meliponini). Sociobiology. 59(3):1-11.

Ferreira-Caliman MJ, Turatti ICC, Lopes NP, Zucchi R, Nascimento FS. 2012. Analysis of insect cuticular compounds by non-lethal solid phase micro extraction with styrene-divinylbenzene copolymers. Journal of Chemical Ecology. 38:418-426.

Ferreira-Caliman MJ, Andrade-Silva ACR, Guidetti-Campos MCG, Turatti ICC, Santos do Nascimento FS, Lopes NP. 2014. A non-lethal SPME method for insect cuticular analysis by GC-MS. *Analytical Methods*. 6:8823.

Fielde AM. 1905. The progressive odor of ants. Biological Bulletin 10(1):1-16.

Finck, Jonas, Emma L. Berdan, Frieder Mayer, Bernhard Ronacher and Sven Geiselhardt. 2016. Divergence of cuticular hydrocarbons in two sympatric grasshopper species and the evolution of fatty acid synthases and elongases across insects. Sci. Rep. **6**, 33695; doi: 10.1038/srep33695.

Gadagkar R. 1985. Kin recognition in social insects and other animals - a review of recent findings and a consideration of their relevance for the theory of kin selection. *Proc. Indian Acad. Sci. (Anim. Sci.)*. 94:587-621.

Gamboa GJ, Reeve HK, Pfennig DW. 1986a. The evolution and ontogeny of nestmate recognition in social wasps. Annual Review of Entomology 31:431-454.

Gamboa GJ, Reeve HK, Ferguson ID, Wacker TL. 1986b. Nestmate recognition in social wasps: the origin and acquisition of recognition odours. Animal Behaviour 34(3):685-695.

Gamboa GJ, Grudzien TA, Espelie KE, Bura EA. 1996. Kin recognition pheromones in social wasps: combining chemical and behavioural evidence. Animal Behaviour 51:625-629.

Gerhardt H, Schmitt C, Betz O, Albert K, Lämmerhofer M. 2015. Contact solid-phase microextraction with uncoated glass and polydimethylsiloxane-coated fibers versus solvent sampling for the determination of hydrocarbons in adhesion secretions of Madagascar hissing cockroaches *Gromphadorrhina portentosa* (Blattodea) by gas chromatography-mass spectrometry. *Journal of Chromatography A*. 1388(3):24-35.

Harrell FE Jr. Package 'Hmisc'. CRAN2018, 235-6. https://cran.r-project.org/package=Hmisc (CRAN, 2019).

Hölldobler, B. and Michener, C. D. (1980). Mechanisms of identification and discrimination in social Hymenoptera. In Evolution of Social Behavior: Hypotheses and Empirical Tests, ed. H. Markl. Weinheim: Chemie Verlag, pp. 35–58.

Jombart T, Devillard S and Balloux, F (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11: 94

Jombart, T. (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24: 1403-1405.

Jombart T and Collins C. 2015. A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0. https://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf

Kingwell C. 2020. The origin of 'queen pheromones': chemical communication as a catalyst for social behavior in the flexibly eusocial bee *Megalopta genalis*. PhD Dissertation, Cornell University, New York.

Kingwell C, Böröczky K, Steitz I, Ayasse M, Wcislo W. 2021. Cuticular and dufour's gland chemistry reflect reproductive and social state in the facultatively eusocial sweat bee *Megalopta genalis* (Hymenoptera: Halictidae). Journal of Chemical Ecology. 47:420-432.

Lacy RC, Sherman PW. 1983. Kin recognition by phenotype matching. The American Naturalist 121(4):489-512.

Lahav S, Soroker V, Hefetz A, Vander Meer RK. 1999. Direct behavioral evidence for hydrocarbons as ant recognition discriminators. Naturwissenschaften. 86:246-249.

Layton JM, Camann MA, Espelie KE. 1994. Cuticular lipid profiles of queens, workers, and males of social wasp *Polistes metricus* Say are colony-specific. Journal of Chemical Ecology 20:2307-2321.

Legan AW, Jernigan CM, Miller SE, Fuchs MF, Sheehan MJ. 2021. Expansion and accelerated evolution of 9-exon odorant receptors in *Polistes* paper wasps. Molecular Biology and Evolution 38(9):3832-3846.

Legan AW, Vogt CC, Sheehan MJ. 2022. Postural analysis reveals persistent vigilance in paper wasps after conspecific challenge. bioRxiv doi: https://doi.org/10.1101/2022.05.25.493496

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

Liaw A, Wiener M. 2002. Classification and regression by randomForest. R News 2(3):18-22.

Liebig J. 2010. Hydrocarbon profiles indicate fertility and dominance status in ant, bee, and wasp colonies. In: Blomquist GJ, Bagnères A-G eds. Insect hydrocarbons: biology, biochemistry, and chemical ecology. Cambridge University Press, Cambridge.

Liu S., Ouyang G. (2017) Introduction of Solid-Phase Microextraction. In: Ouyang G., Jiang R. (eds) Solid Phase Microextraction. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-53598-1_1

Loope KJ. 2015. Chapter 6: Social organization predicts cuticular hydrocarbon diversity in paper wasps. In “Reproductive conflicts and signal evolution in social wasps and bees.” PhD Dissertation, Cornell University, New York. p. 153-190.

Lorenzi MC, Bagnères AG, Clèment J-L, Turillazzi S. 1997. *Polistes biglumis bimaculatus* epicuticular hydrocarbons and nestmate recognition (Hymenoptera, Vespidae). Insectes Sociaux 44:123-138.

Martin S, Helanterä H, Drijfhout FP. 2008a. Colony-specific hydrocarbons identify nest mates in two species of *Formica* ant. *J. Chem. Ecol*. 34:1072-1080.

Martin SJ, Vitikainen E, Helanterä H, Drijfhout FP. 2008b. Chemical basis of nest-mate discrimination in the ant *Formica exsecta*. Proceedings of the Royal Society B.

Martin S and Drijfhout F. 2009. A review of ant cuticular hydrocarbons. *Journal of Chemical Ecology*, 35, 1151 doi:10.1007/s10886-009-9695-4

Martin SJ, Zhong W, Drijfhout FP. 2009. Long-term stability of hornet cuticular hydrocarbons facilitates chemotaxonomy using museum specimens. *Biological Journal of the Linnean Society*. 96(4):732-737.

Monnin T, Malosse C, Peeters C. 1998. Solid-phase microextraction and cuticular hydrocarbon differences related to reproductive activity in queenless ant *Dinoponera quadriceps*. *Journal of Chemical Ecology*. 24(3):473-490.

Murakami ASN, Nunes TM, Desuó IC, Shima SN, Mateus S. 2015. The cuticular hydrocarbons profiles in the colonial recognition of the neotropical eusocial wasp, *Mischocyttarus cassununga* (Hymenoptera: Vespidae). *Sociobiology*. 62(1): 109-115.

Neupert S, Hornung M, Millar JG, Kleineidam CJ. 2018. Learning distinct chemical labels of nestmates in ants. Frontiers in Behavioral Neuroscience. 12:191.

Niehuis O, Buellesbach J, Gibson JD, Pothmann D, Hanner C, Mutti NS, Judson AK, Gadau J, Ruther J, Schmitt T. 2013. Behavioural and genetic analyses of *Nasonia* shed light on the evolution of sex pheromones. Nature. 494:345-348.

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.

Oksanen J, Simpson GL, Blanchet G, Kindt R, Legendre P, Minchin PR, O'Hara RB, Solymos P, Stevens MHH, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista HBA, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill MO, Lahti L, McGlinn D, Oullette M-H, Cunha ER, Smith T, Stier A, Braak CJFT, Weedon J. 2022. vegan: Community Ecology Package. R package version 2.6-2. https://CRAN.R-project.org/package=vegan

Ortiz CC, Tibbetts EA. 2020. Visual and chemical signals provide different information in *Polistes fuscatus* wasps. Ethology. 127:231-237.

Otte T, Hilker M, Geiselhardt S. 2018. Phenotypic plasticity of cuticular hydrocarbon profiles in insects. *J Chem Ecol*. 44:235-247.

Ottensmann M, Stoffel MA, Nichols HJ, Hoffman JI. GCalignR: an R package for aligning gas-chromatography data for ecological and evolutionary studies. PLOS ONE. 13(6):e0198311.

Page RE, Metcalf RA, Metcalf RL, Erickson EH, Lampman RL. 1991. Extractable hydrocarbons and kin recognition in honeybee (*Apis mellifera* L.). Journal of Chemical Ecology. 17(4):745-756.

Pask GM, Slone JD, Millar JG, Das P, Moreira JA, Zhou X, Bello J, Berger SL, Bonasio R, Desplan C, Reinberg D, Liebig J, Zwiebel LJ, Ray A. 2017. Specialized odorant receptors in social insects that detect cuticular hydrocarbon cues and candidate pheromones. *Nature Communications* 8:297.

Pfennig DW, Reeve HK, and Shellman JS. 1983. Learned component of nestmate discrimination in workers of a social wasp, *Polistes fuscatus* (Hymenoptera: Vespidae). *Anim. Behav.* 31:412-416.

Pfennig DW, Gamboa GJ, Reeve HK, Reeve JS, and Ferguson ID. 1983. The mechanism of nestmate discrimination in social wasps (*Polistes*, Hymenoptera: Vespidae). *Behav. Ecol. Sociobiol.* 13: 299-305.

Pickett KM, McHenry A, Wenzel JW. 2000. Nestmate recognition in the absence of a pheromone. Insectes soc. 47:212-219.

Post DC, and Jeanne RL. 1982. Recognition of former nestmates during colony founding by the social wasp *Polistes fuscatus* (Hymenoptera: Vespidae). *Behav. Ecol. Sociobiol.* 11: 283-285.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/.

Reeve HK. (1989). The evolution of conspecific acceptance thresholds. *The American Naturalist*. 133:407-435.

Ryan RE, Forbes GC, Gamboa GJ. 1984. Male social wasps fail to recognize their brother (*Polistes fuscatus*, Hymenoptera: Vespidae). Journal of the Kansas Entomological Society. 57(1):105-110.

Ryan, Ruth E., and George J. Gamboa. 1986. Nestmate recognition between males and gynes of the social wasp *Polistes fuscatus* (Hymenoptera: Vespidae). *Ann. Entomol. Soc. Am.* 79: 572-575.

Shellman, Janet S., and George J. Gamboa. 1982. Nestmate discrimination in social wasps: the role of exposure to nest and nestmates (*Polistes fuscatus*, Hymenoptera: Vespidae). *Behavioral Ecology and Sociobiology*. 11: 51-53.

Sims C, Birkett MA, Withall DM. 2022. Enantiomeric discrimination in insects: the role of OBPs and ORs. Insects. 13(4):368.

Singer TL, Espelie KE. 1992. Social wasps use nest paper hydrocarbons for nestmate recognition. Animal Behaviour. 44:63-68.

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. *Journal of Chemical Ecology* 18:785-797.

Singer TL, Espelie KE. 1996. Nest surface hydrocarbons facilitate nestmate recognition for the social wasp, *Polistes metricus* Say (Hymenoptera: Vespidae). Journal of Insect Behavior 9(6):857-870.

Singer TL. 1998. Roles of hydrocarbons in the recognition systems of insects. American Zoologist. 38:394-405.

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

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.

Starks PT, Fischer DJ, Watson RE, Melikian GL, Nath SD. (1998). Context-dependent nestmate discrimination in the paper wasp, *Polistes dominulus*: a critical test of the optimal acceptance threshold model. *Animal Behaviour* 56, 449-458.

Tibbetts EA, Liu M, Laub EC, Shen S-F. 2020. Complex signals alter recognition accuracy and conspecific acceptance thresholds. Phil. Trans. R. Soc. B 375:20190482.

Tsutsui ND. 2004. Scents of self: the expression component of self/non-self recognition systems. Ann. Zool. Fennici 41:713-727.

Turillazzi S, Sledge MF, Moneti G. 1998. Use of a simple method for sampling cuticular hydrocarbons from live social wasps. Ethology Ecology and Evolution 10:293-297.

Van Oystaeyen A, Oliveira RC, Holman L, van Zweden JS, Romero C, Oi CA, d'Ettorre P, Khalesi M, Billen J, Wäckers F, Millar JG, Wenseleers T. 2014. Conserved class of queen pheromones stops social insect workers from reproducing. Science. 343:287-290.

Van Wilgenburg E, Felden A, Choe D-H, Sulc R, Luo J, Shea KJ, Elgar MA, Tsutsui ND. 2012. Learning and discrimination of cuticular hydrocarbons in a social insect. Biology Letters. 8:17-20.

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.

Van Zweden JS, Dreier S, d'Ettorre P. 2009. Disentangling environmental and heritable nestmate recognition cues in a carpenter ant. Journal of Insect Physiology. 55:158-163.

Vernier C.L., J.L. Krupp, K. Marcus, A. Hefetz, J.D. Levine, and Y. Ben-Shahar. 2019. The cuticular hydrocarbon profiles of honey bee workers develop via a socially-modulated innate process. *eLife* 8:e41855.

Walsh, J., L. Pontieri, P. d'Ettorre, and T.A. Linksvayer. 2020. Ant cuticular hydrocarbons are heritable and associated with variation in colony productivity. *Proc. R. Soc. B.* 287:20201029

Wickham H. ggplot2: elegant graphics for data analysis. Springer-Verlag New York, 2016.

Wyatt TD. 2010. Pheromones and signature mixtures: defining species-wide signals and variable cues for identity in both invertebrates and vertebrates. Journal of Comparative Physiology A. 196:685-700.

CHAPTER 2

EXPANSION AND ACCELERATED EVOLUTION OF 9-EXON ODORANT RECEPTORS IN *POLISTES* PAPER WASPS

Andrew W. Legan, Christopher M. Jernigan, Sara E. Miller, Matthieu Fuchs, Michael Sheehan.

Published in *Molecular Biology and Evolution* Volume 38, Issue 9, Pages 3832–3846.

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.

Background

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 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 TB 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. 2.1). 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.

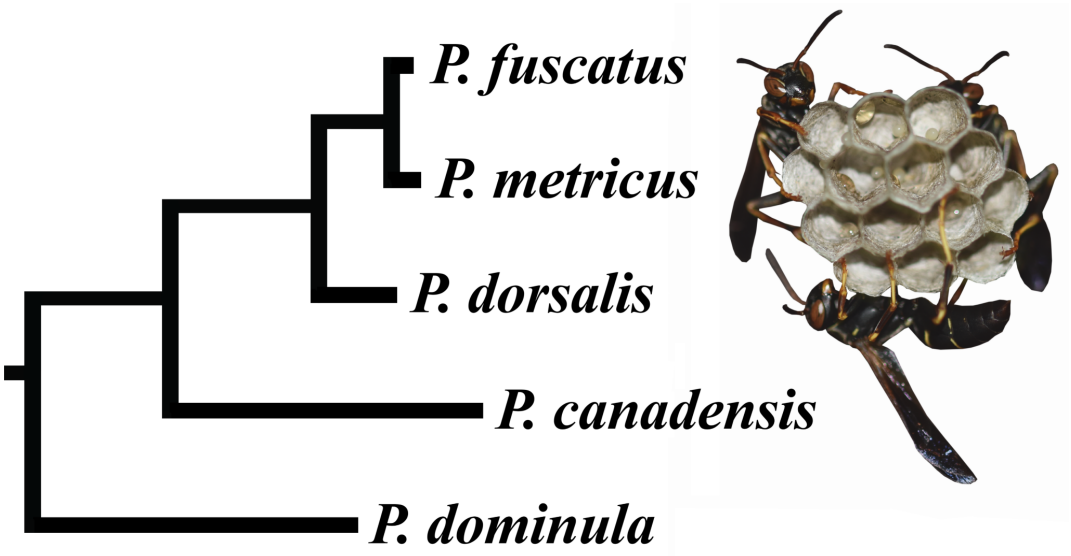

Figure 2.1. Phylogeny of five *Polistes* species considered in this study: *P. fuscatus*, *P. metricus*, *P. dorsalis*, *P. canadensis*, and *P. dominula*. The photo to the right of the phylogeny shows *P. fuscatus* foundresses on a nest. Phylogenetic tree based on the 16S ribosomal RNA gene and the cytochrome oxidase subunit I gene.

Methods

Antennal Lobe Imaging

Immunocytochemistry:

In the fall of 2018, 20 adult *P. fuscatus* wasps (10 male and 10 gyne) were collected in Ithaca, NY. Wasps were immobilized by cooling at 4°C, heads were removed, and brains were dissected from the head capsule and placed in ~1 ml of 4% paraformaldehyde in 0.01M phosphate buffered saline (PFA) overnight at 4°C. Mouse monoclonal anti-synapsin antibodies (SYNORF1, 3C11; Data Bank Hybridoma) were used as primary antibodies to label synaptically rich antennal lobe glomeruli. Phalloidin conjugated with TRITC (Tetramethylrhodamine Isothiocyanate; Invitrogen) was used to distinguish glomeruli borders and label antennal nerve tracts. F(ab')2 fragments of donkey anti-mouse antibodies conjugated to Alexa 488 were used as secondary antibodies to visualize anti-synapsin (Jackson ImmunoResearch Laboratories). After overnight fixation, brains were washed in 0.01M phosphate buffered saline with Triton-X 100 (PBS-TX) 6 times for 20 min. Brains were incubated for three nights at room temperature with anti-synapsin at 1:800 in PBS-TX solution. Brains were then washed again in PBS-TX (6 times for 20 min) and incubated for three nights at RT with donkey anti-mouse antibodies conjugated with Alexa 488 at 1:270, and then with Phalloidin-TRITC at 1:160 in PBS-TX. After staining with secondary antibodies, brains were post-fixed with 4% PFA for 10 min and dehydrated using increasing concentrations of methanol. Following full dehydration and two washes in 100% methanol, brains were cleared in methyl salicylate overnight. They were then mounted on slides in methyl salicylate.

Imaging:

Images were collected using an inverted 880 confocal laser scanning microscope (Zeiss LSM880 Indimo, AxioObserver) with a C-Apochromat10x/0.45 water-immersion objective and a 1.8x digital zoom using the appropriate filter and laser setting for each fluorescent molecule and 8x line averaging. Image stacks were collected using 10-15 μm optical sections for both males and gynes. To individually label the glomeruli from confocal image stacks we used AVIZO and Amira software (Thermo Scientific, FEI) to make 3D reconstructions of the antennal lobe from whole mount preparations. Confocal stack files were imported into AVIZO software, and the voxel dimension outputs of each image stack were imported to provide correct dimensions (Thermo Scientific, FEI). Using the image segmentation function, we then identified glomeruli by hand throughout the entire image stack.

Antennal RNAseq

P. fuscatus mRNA library preparation and sequencing:

In the fall of 2017, 12 adult *P. fuscatus* wasps (6 males and 6 gynes) from Ithaca, NY and Watkins Glen, NY were freeze-killed on dry ice and stored at -80°C. Each pair of antennae was cut at the base of the scape and weighed (paired gyne antennae = 1.27 mg ± 0.26 SD; paired male antennae = 1.57 mg ± 0.44 SD). RNA was isolated from 2 pooled antennae of each individual using the PureLink RNA Micro Kit (Ambion). Libraries were prepared using the NEBNext Ultra II RNA Library Prep Kit for Illumina (NEB #E7770L) with the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB #E7490) and NEBNext Multiplex Oligos for Illumina (NEB #E7600S). Libraries were analyzed on a fragment analyzer for quality

control and sequenced on two lanes of an Illumina HiSeq (Illumina, San Diego, CA, USA).

Mapping mRNA reads of *P. fuscatus* and *P. dominula*:

Raw reads from *P. fuscatus* antennae were trimmed using Trimmomatic v.0.39 to (1) remove adapter sequences, (2) cut leading and trailing bases below quality of 3, (3) cut 4-base wide windows when the average quality per base dropped below 20, and (4) remove reads below 35 base pairs (bp) long (Bolger et al. 2014). After trimming and filtering, more than 154 million paired end reads remained (75,846,826 male; 78,453,620 gyne) with an average length of 132 bp. Processed paired end reads were mapped, first to the *P. fuscatus* genome and subsequently to the genomes of two other *Polistes* species in the fuscopolistes clade (*P. metricus* and *P. dorsalis*), using STAR v2.6 to generate BAM files sorted by genome coordinate (Dobin et al. 2013). Each read was allowed to map to no more than one locus. Mapped reads from each sex were combined and assembled using Trinity genome-guided assembly (Haas et al. 2013). Summary statistics for the assembly are shown in table 2.1. Raw reads from the whole-body transcriptome of a single individual *P. dominula* were downloaded from the Sequence Read Archive (Run SRR4301488; Lopez-Osorio et al. 2017). To aid in manual annotation of ORs, these reads were aligned to the *P. dominula* genome with STAR v2.7, using the same parameters as described above for *P. fuscatus*.

Table 2.1. Summary statistics of antennal transcriptome assembly

Total trinity 'genes'	163537		
Total trinity transcripts	211491		
Percent GC	33.41	Based on longest isoform per 'gene'	
contig N10	13311	contig N10	7460
contig N20	8798	contig N20	4062
contig N30	6356	contig N30	2259
contig N40	4547	contig N40	1383
contig N50	3026	contig N50	953
Median contig length	463	Median contig length	399
Average contig	1152.76	Average contig	703.89
Total assembled bases	243799125	Total assembled bases	115111742

Gene Annotation

OR annotation in *P. fuscatus*:

The *Polistes fuscatus* genome and annotations were accessed through NCBI (Miller SE et al. 2020). Genome scaffolds containing putative ORs were identified based on sequence similarity with published insect ORs from a range of species: *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, Robertson et al. 2010, Zhou et al. 2012, Zhou et al.

2015, Brand and Ramirez 2017, McKenzie and Kronauer 2018). The *P. fuscatus* genome was queried with amino acid sequences from each of the listed species using the TBLASTN algorithm with an e-value cutoff of 1e-5 (Altschul et al. 1997). Putative coding regions were identified by examining TBLASTN match regions and chaining the match regions according to criteria similar to those used by Zhou et al. (2012): neighboring match regions were included in the same coding region if the match regions were in close proximity and if the upstream amino acid query sequence was N-terminal to that of the downstream match region. Coding sequence (CDS) was determined by manually curating exon-intron boundaries in Geneious v11.1.5 using evidence from MAKER automated predictions, TBLASTN homology, Trinity predicted transcripts (aligned to genome using BLAT) and mapped RNAseq reads. We consulted version 24 of the Havana annotation guidelines to inform manual annotation of exon-intron splice sites (March 2016 guidelines by the Human and Vertebrate Analysis and Annotation group at WTSI). "AG" was the only permitted acceptor site. "GT" and "GC" were both acceptable donor sites, although the "GC" donor was only present in a small number of genes. Amino acid sequences of ORs generated by this pipeline were used as queries in an iterative round of TBLASTN with scoring matrix PAM30 (Pearson 2013), to search for species-specific expansions, especially in the highly divergent 9-exon receptor clade. 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. fig. 2.2).

Figure 2.2. Screenshot of Geneious genome browser view of a portion of 9-exon tandem array s13 highlighting the usefulness of antennal RNAseq reads and the necessity to manually modify automated gene predictions during OR gene annotation. The blue coverage map represents the RNAseq reads, aligned on the bottom track in the inset. The top annotation track shows manually annotated OR genes (green = putatively functional; orange = pseudogenes) and the bottom annotation track shows the automated annotations generated by the MAKER pipeline, which was otherwise highly effective in annotating most genes (BUSCO 97.7% eukaryotic genes, table 2.2).

OR annotation in 4 other *Polistes*:

The genomes and annotations for *Polistes canadensis*, *Polistes dominula*, *Polistes dorsalis*, and *Polistes metricus* were accessed through NCBI (Patalano et al. 2015; Standage et al. 2016; Miller SE et al. 2020). Putative ORs were identified by using TBLASTN with the same sample of Hymenopteran protein query sequences used during manual annotation of *P. fuscatus* ORs, with the addition of *P. fuscatus* OR protein sequences. Coding sequence (CDS) was determined by manually curating exon-intron boundaries in Geneious v11.1.5, using evidence from automated predictions and TBLASTN homology. Antennal RNAseq reads from *P. fuscatus* were

mapped to the *P. metricus* and *P. dorsalis* genomes to aid in manual OR annotation in these species. Adult whole-body RNAseq reads from *P. dominula* were mapped to the *P. dominula* genome to inform manual gene annotation in this species. Uncertain gene models were aligned using Muscle version 3.8.425 with maximum 4 iterations and gene models were manually adjusted based on homology with putatively functional ORs in *P. fuscatus* (Edgar 2004). Gene models were called pseudogenes if they exhibited frame-shift mutations, premature stop codons, or unacceptable 5' donor or 3' receptor splice sites.

Naming ORs:

All ORs were named with a species-specific prefix using the first letter of the genus followed by the first three letters of the species epithet (e.g. PfusOR#). Orco was not numbered (e.g. PfusOrco), and no OR was labeled OR1, since some previous studies refer to Orco as OR1. All tuning ORs were numbered in the order in which they appeared in the genome, beginning with OR2. Therefore, OR names with the same number do not necessarily indicate orthology among *Polistes*. Incomplete gene models were named with a suffix to indicate which features of the gene model are missing, following Oxley et al. (2014): "NTE" missing N-terminus, "INT" missing interior sequence, "CTE" missing C terminus, "NI" missing N terminus and interior sequence, "NC" missing N and C termini. All pseudogenes were given the suffix "PSE."

Search for ORs in unassembled genomic reads:

Genome misassembly is a potential cause of overestimation of pseudogenes, and is more common in repetitive regions of the genome. For example, many ORs

that were previously annotated as pseudogenes in the clonal raider ant were not identified in a new, chromosome arm level assembly of the genome, which also yielded many more functional OR gene models (McKenzie and Kronauer 2018). Given the completeness of the *Polistes* genomes (table 2.2) and the close correspondence between functional *P. fuscatus* ORs and number of *P. fuscatus* antennal lobe glomeruli, this type of error likely accounts for minimal annotation errors in this species. We searched for sequence similarity to ORs in the unassembled genomic reads of *P. dorsalis*, the species in which we initially found the fewest functional ORs. The 10x Genomics Supernova Assembler version 2.1.1 failed to assemble unmapped *P. dorsalis* genomic reads greater than or equal to 150 nucleotides. Therefore, the reads themselves were formatted as a blast database and queried using TBLASTN at an e-value cutoff of 1e-5. The same sample of Hymenopteran protein sequences used during manual annotation, with the addition of the putatively functional OR protein sequences of all five *Polistes* species, was used as query during the TBLASTN search of the unaligned *P. dorsalis* genomic reads. This resulted in significant hits to 123 paired-end reads. Given that on average 159 ± 555 SD paired-end reads map to a given OR gene in *P. fuscatus*, it is unlikely that many ORs were lost in unassembled genomic reads. However, it cannot be ruled out that genome misassembly may account for overestimation of pseudogenes in some cases.

Table 2.2. Summary statistics for *Polistes* genome assemblies and annotations

Species	Size[a]	N50[b]	Genes[c]	BUSCO euk.[d]	CEGMA euk.[e]	Reference[f]
P. fuscatus	219	9.11	12034	97.7% (296/303)	N/A	Miller et al. 2020 (GCA_010416935.1)
P. metricus	220	4.63	12032	98.0% (297/303)	N/A	Miller et al. 2020 (GCA_010416925.1)
P. dorsalis	209	5.37	11311	98.0% (297/303)	N/A	Miller et al. 2020 (GCA_010416905.1)
P. canadensis	211	0.52	16997	N/A	98.8% (245/248)	Patalano et al. 2015 (GCA_001313835.1)
P. dominula	208	1.63	12153	N/A	99.2% (246/248)	Standage et al. 2016 (GCA_001465965.1)

a. Size in Mb of genome assembly
b. Scaffold N50 in Mb
c. Total number of annotated genes in genome
d. Annotation completeness as measured by BUSCO score for core eukaryotic genes
e. Annotation completeness as measured by CEGMA score for core eukaryotic genes
f. Reference to genome publication followed by the GenBank assembly accession

Quality assessment of putatively functional ORs:

All *Polistes* OR protein sequences greater than 300 amino acids in length were analyzed using InterProScan version 5.17-56.0 to check for the InterPro olfactory receptor annotation IPR004117, the 7tm odorant receptor Pfam accession PF02949, and the odorant receptor PANTHER accession PTHR21137 (Jones et al. 2014). For each of 192 orthologous groups predicted by OrthoFinder, the orthologs were aligned using Muscle version 3.8.425 with maximum 4 iterations, and alignments were inspected for anomalies that could be the result of annotation error. Adjustments were made to the gene models in 29 orthologous groups. The majority of these adjustments

were changes in the lengths of already annotated exons with few or no mapped mRNA reads.

Hymenoptera-wide analysis of OR gene family:

We used published OR sequences from *Apis mellifera*, *Nasonia vitripennis* (Robertson and Wanner 2006; Robertson et al. 2010), and *Camponotus floridanus* (Zhou et al. 2012), along with OR sequences of *P. fuscatus* (this study) to reconstruct the phylogeny of the OR gene family. All amino acid (aa) sequences of ORs at least 300 aa in length were aligned using MAFFT version 7.453 with the high accuracy "E-INS-i" method ("genafpair" option combined with 1000 cycles of iterative refinement ("maxiterate")) (Katoh and Standley 2013). The alignment was trimmed using trimAl v3 with the "automated1" option to automatically choose the parameters that best fit the qualities of the alignment (Capella- Gutierrez et al. 2009), generating a phylip file with 1066 sequences of 231 aa length. Gene trees were generated with RAxML version 8.2.12 using the Maximum Likelihood optimality criterion with 100 rapid bootstrap inferences, CAT model of rate heterogeneity across sites (Stamatakis 2006), JTT amino acid replacement matrix, F option to use empirical base frequencies from the input alignment ("-m PROTCATJTTF"), "-p 12345" to set initial random number for parsimony search, *N. vitripennis* Orco as outgroup, "-f a" to combine bootstrap and maximum likelihood searches in one run, "-x 12345" to set initial random number for bootstrap search, "-N 100" to perform a bootstrap analysis informed by 100 unique maximum likelihood trees, and "-k" to add branch lengths to trees (Stamatakis 2014, Jones DT et al. 1992). The tree was rerooted in FigTree v1.4.4 to visually clarify the position of the Orco gene clade (available online at

https://github.com/rambaut/figtree/releases/tag/v1.4.4). The tree was visualized in R version 3.6.1 using ggtree (Yu et al. 2017; R Core Team 2019).

Estimation of gene gain and loss events:

Gene duplication and loss events were reconstructed by reconciling a gene tree with a species tree in NOTUNG version 2.9.1.3 (Chen et al. 2000; Durand et al. 2006; Vernot et al. 2008). A gene tree of putatively functional OR proteins at least 300 amino acids in length was generated using published protein sequences for 4 ants (*A. echinatior*, *A. cephalotes*, *S. invicta*, *C. floridanus*; Zhou et al. 2012, 2015), 2 bees (*A. mellifera*, *L. albipes*; Robertson et al. 2010, Zhou et al. 2015), 3 nonsocial wasps (*N. vitripennis*, *M. demolitor*, *C. solmsi*; Robertson et al. 2010, Zhou et al. 2015), and 5 social wasps (*P. fuscatus*, *P. metricus*, *P. dorsalis*, *P. canadensis*, *P. dominula*; annotated in this study). The gene tree was generated as described above for the Hymenoptera OR tree, without rerooting. Following Zhou et al. 2012, all internal branches with bootstrap support values below 70 were collapsed into polytomies in R using ggtree. Notung was run in rearrangement mode using an edge weight (bootstrap value) threshold of 90, with cost of "1.5" for a duplication and cost of "1" for a loss.

Analysis of *Polistes* OR gene family:

The OR sequences of *P. fuscatus*, *P. metricus*, *P. dorsalis*, *P. canadensis*, and *P. dominula* were filtered to include only ORs of at least 300 amino acids in length. These were aligned using MAFFT ("E-INS-i" method) and trimmed in trimal using the parameters described above for the Hymenoptera OR tree. The tree was generated using RAxML, rerooted in FigTree, and visualized in R using ggtree as described above. Trees in the supplementary materials were visualized using Interactive Tree Of

Life (iTOL) version 5.6.2 (Letunic and Bork 2019). Orthology of *Polistes* OR amino acid sequences was inferred using sequence similarity and gene microsynteny in the genome. Orthologs were called based on similarity of amino acid sequences with OrthoFinder version 2.2.7 using DIAMOND for sequence similarity searches with an e-value cutoff of 0.001 (Emms and Kelly 2015). OrthoFinder predicted a total of 217 orthogroups, 135 of which included gene copies from both *P. fuscatus* and *P. dorsalis*. For the purpose of pairwise d_N/d_S analysis, this set of orthologs was expanded by inspecting the gene tree and Muscle alignments (see "Genomic organization" below) of syntenic regions. Gene copies within a clade supported by a bootstrap value of at least 70 and located collinearly in microsyntenic regions of the genome were considered orthologs in the expanded ortholog set of 150 single copy orthologs between *P. fuscatus* and *P. dorsalis*.

Polistes species phylogeny:

The phylogenetic tree of *Polistes* species in fig. 2.1 was built from Supplemental Fig. 2 in Sheehan et al. 2015 using the "physketch" function in the phytools R package (Revell 2012). The initial tree was constructed using the 16S ribosomal RNA gene and the cytochrome oxidase subunit I gene (Sheehan et al. 2015).

Genomic Organization:

Two functional ORs separated by a pseudogenized OR were considered neighbors if their coding sequences were within 5kb of each other. Neighboring OR protein sequences were aligned using Muscle v3.8.425 with maximum 4 iterations to determine percent amino acid identity. One-way ANOVA comparing percent amino

acid identity of neighbors across eight tandem arrays was performed using the "aov" function, and Tukey HSD test with the "TukeyHSD" function, both of which are in the base R package stats version 3.6.1 (R Core Team 2019).

Sequence Analyses:

Tests for positive selection: All putatively functional *Polistes* ORs greater than 350 amino acids in length were used in analyses of episodic diversifying positive selection within OR subfamilies using aBSREL in HyPhy version 2.5.15 (Smith MD et al. 2015; Pond et al. 2020). Subfamily- specific alignments were initially generated using MAFFT version 7.453 ("E-INS-i") and trimmed using trimal with the "gappyout" option. Maximum-likelihood trees were generated using RAxML as above, except the "PROTGAMMAAUTO" model was used to automatically choose the best protein model given likelihood. Codon alignments were generated using trimal with the "backtranslate" and "gappyout" options. The significance threshold for each aBSREL analysis was Holm-Bonferroni corrected $p = 0.05$.

Estimation of d_N/d_S:

Values of pairwise d_N/d_S for orthologs shared by *P. fuscatus* and *P. dorsalis* were calculated using PAML version 4.9 (Yang 2007) program yn00 with the Yang and Nielsen (2000) method. Values of *d*N, *d*S, and omega were compared between 9-exon and non-9- exon orthologous pairs using the function "t.test" in the base R package stats version 3.6.1. Next, amino acid sequences of ORs in 145 orthogroups predicted by OrthoFinder were independently aligned, alignments trimmed in trimal with the "-gappyout" option, and ML trees generated using RAxML as above for subfamily-level selection analyses. This analysis was limited to orthogroups

containing at least four genes, including inparalogs. Codon alignments were generated as above. PAML CodeML model M0 was used to estimate values of dN/dS for each orthogroup. Control file parameters for model M0 were identical to those in supplementary table S3 of Gao et al. (2019).

Test for different intensities of selection:

RELAX analysis version 3.1.1 was run in HyPhy version 2.5.15 to test for relaxed selection within the 9-exon OR subfamily relative to all other OR subfamilies (Wertheim et al. 2015; Pond et al. 2020). Alignment, trimming, backtranslation, and phylogenetic reconstruction of all putatively functional *Polistes* OR amino acid sequences greater than 350 aa in length was carried out using the same programs and parameters as above for analysis of positive selection. Ambiguous nucleotides in OR CDS were replaced by the character “N” before backtranslation to generate the codon alignment. The 9-exon OR clade was set as the “test” set of branches, and all other branches in the tree constituted the “reference” set. Foreground branches were labeled using the HyPhy phylotree widget (http://phylotree.hyphy.org/). The RELAX output file is available in the online supplement in json format.

Test for gene conversion:

To search for gene conversion events, all *Polistes* OR subfamily codon alignments were analyzed using GENECONV version 1.81a with options "/lp" to list all pairwise significant fragments, "/w123" to set initial random number, and "-nolog" to avoid a segmentation fault in UNIX (Sawyer 1989). Bonferroni-corrected KA p-values below 0.05 were considered significant.

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) (fig. 2.3). Across a sample of insects, the number of 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. 2.4). 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) (table 2.2). 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).

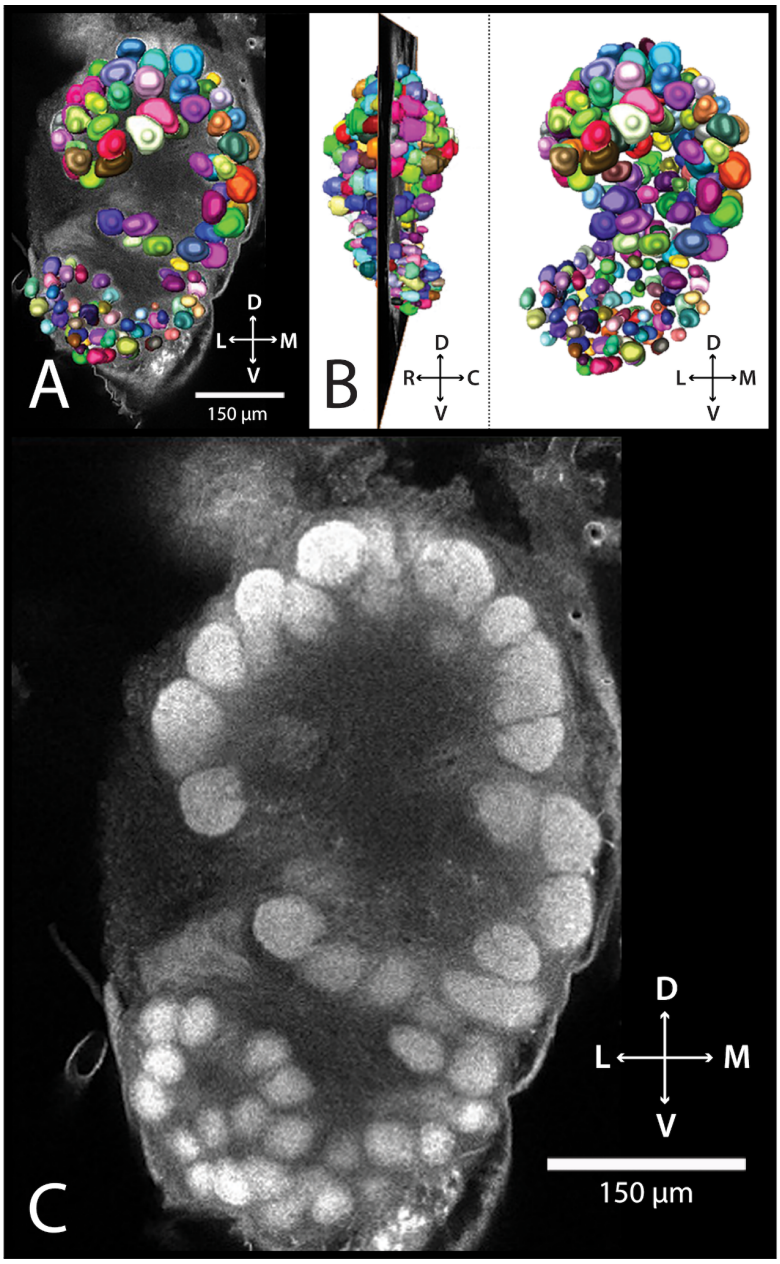

Figure 2.3. (A) 3D reconstruction of the anti-synapsin stained left antennal lobe of a P. fuscatus gyne superimposed on an orthogonal slice of a confocal image stack. (B) Sagittal view of 3D reconstruction with orthogonal slice corresponding to (A), accompanied by frontal view on the right. Colors denote individual glomeruli. (C) Confocal microscopy slice. Scale bars represent 150 µm. D: dorsal; V: ventral; M: medial; L: lateral; R: rostral; C: caudal.

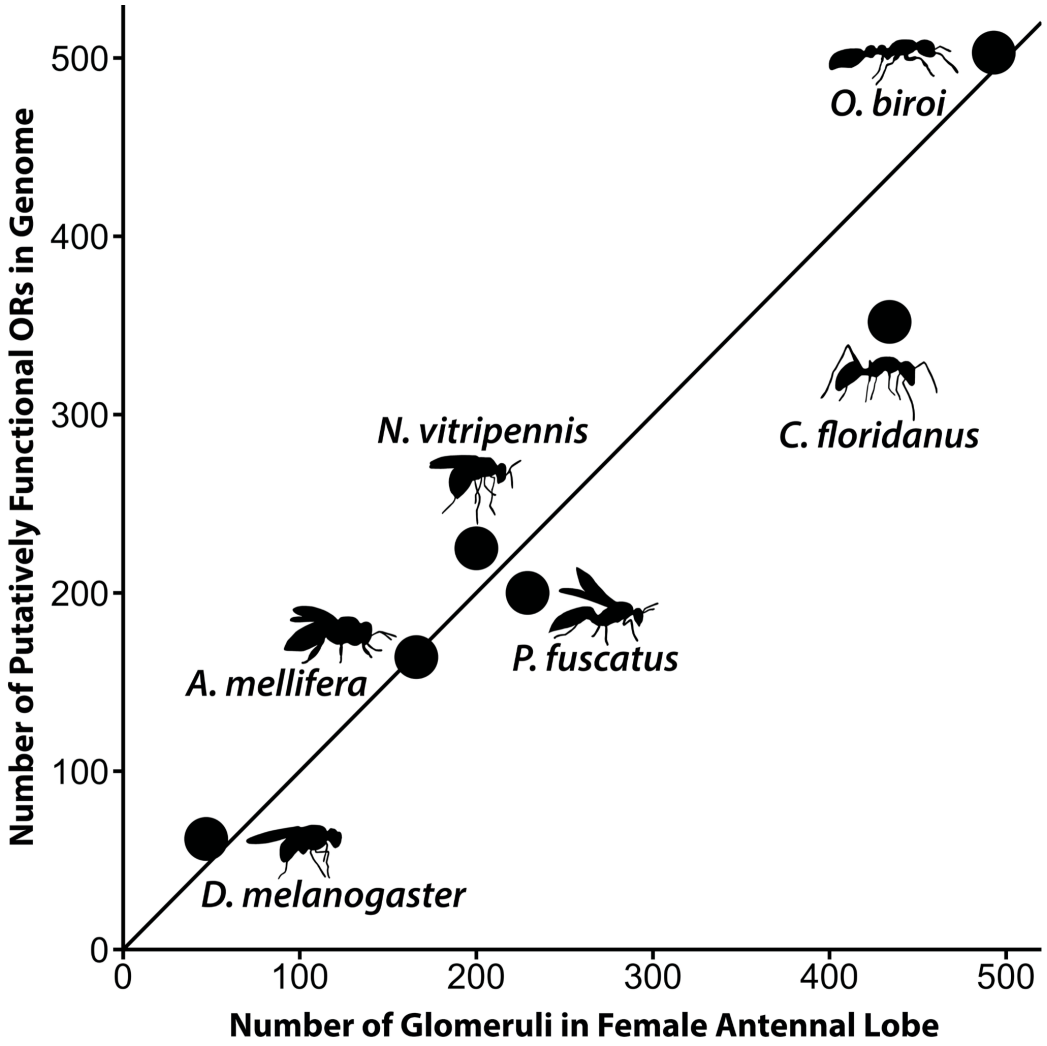

Figure 2.4. The number of functional ORs is correlated with the number of antennal lobe glomeruli across insect species (50 glomeruli and 62 ORs in the genome of the common fruit fly *D. melanogaster*, Fishilevich and Vosshall 2005; 166 glomeruli in the worker and 163 intact ORs in the genome of the 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.

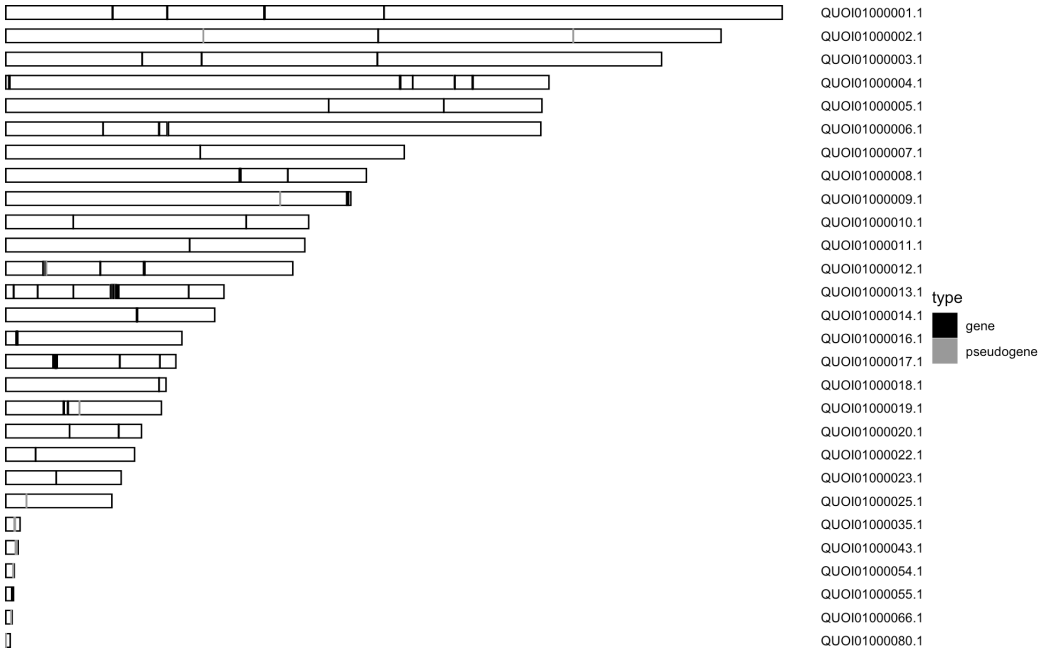

Figure 2.5. Ideograms of 28 *Polistes fuscatus* reference genome scaffolds which contain OR genes (black) and OR pseudogenes (gray) represented by vertical lines. The last unique digits of each scaffold name correspond to the scaffold numbers (*s#*) referred to in the main text (e.g. "QUOI01000017.1" corresponds to the main text scaffold "*s17*").

Manual gene annotation of P. fuscatus ORs recovered 231 gene models across 28 scaffolds (fig. 2.5), 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 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 2.3). 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 (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 2.3.

Table 2.3. Odorant receptor subfamily size and orthology across *Polistes*

Subfam	*P. fus.*	*P. met.*	*P. dor.*	*P. can.*	*P. dom.*	proport. 1:1 ortho.[a]	ortho. ident.[b]
9-exon	105	109	85	101	92	32	91 ± 4.8
A	1	1	1	1	1	100	85.6
B	1	1	1	1	1	100	95.9
C	1	1	1	1	1	100	96.6
D	1	1	1	0	1	0	NA
E	8	7	6	6	5	63	94.5 ± 2.7
F	5	5	5	4	5	80	96.6 ± 1.5
G	2	2	2	1	2	0	NA
H	9	8	5	6	6	33	93.3 ± 0.5
I	1	1	1	1	1	100	95.2
J	4	4	4	4	4	100	96.2 ± 1.4
K	1	1	1	1	1	100	98.4
L	25	25	25	23	22	60	94.9 ± 2.7
M	1	1	1	1	1	100	92.5
O	3	3	3	3	4	67	92.1 ± 2.4
P	1	1	1	1	1	100	94.9
Q	5	7	8	7	6	80	94.7 ± 1.5
R	1	1	1	1	1	100	95.3
T	14	14	14	14	14	71	94.91 ± 3.1
V	6	6	6	6	6	100	96.2 ± 1.2
W	1	1	1	1	1	100	92
X	1	1	1	1	1	100	97.3
Z	1	1	1	1	1	100	95.5
Orco	1	1	1	1	1	100	99.9
Unclass.	1	1	1	1	1	100	92.3
All	200	204	177	188	180	50	93.6 ± 4

a. proportion of *P. fuscatus* ORs with 1:1 orthology across all species
b. mean pairwise amino acid identity among OrthoFinder 1:1 orthologs (%) ± SD

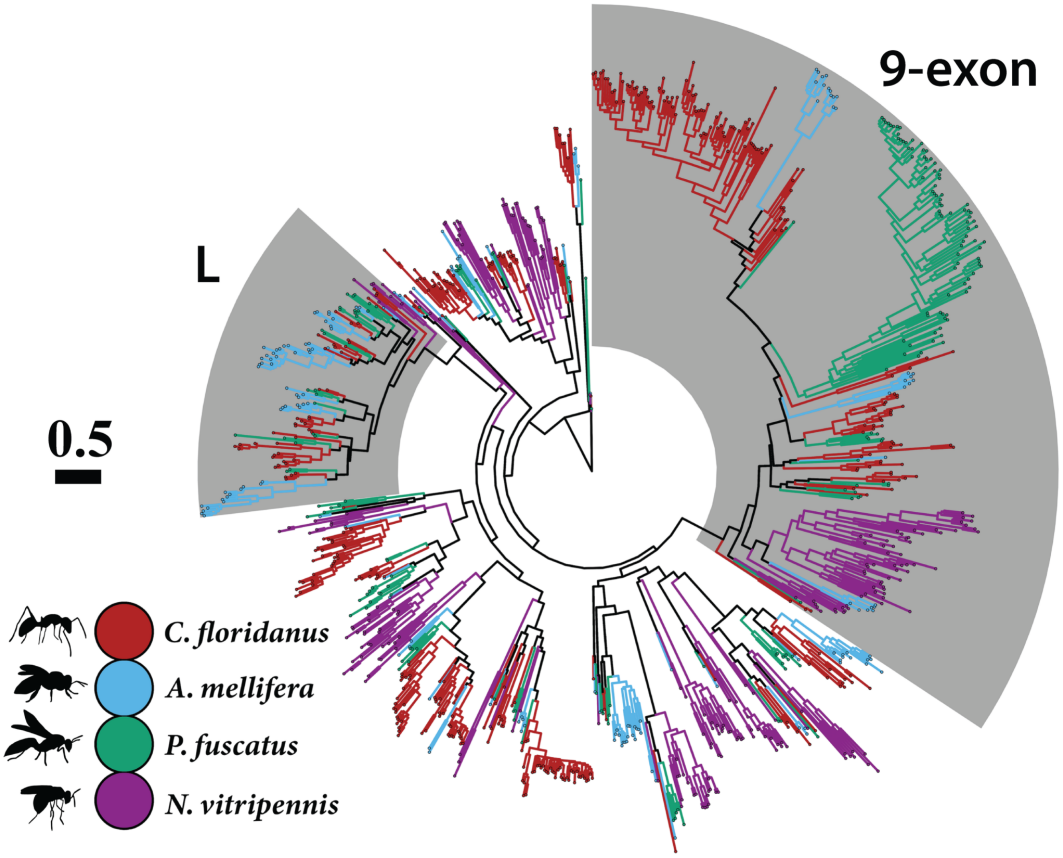

Figure 2.6. Maximum likelihood OR protein tree constructed using data from four Hymenopterans (*Apis mellifera*, Robertson and Wanner 2006; *Camponotus floridanus*, Zhou et al. 2012; *Nasonia vitripennis*, Robertson et al. 2010). Branches are colored by species (Red: *C. floridanus*; Light blue: *A. 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.

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 (fig. 2.6). Gene gain and loss events were predicted using NOTUNG (Chen et al. 2000) and mapped onto a species cladogram of 14 Hymenopterans (fig.

2.7). 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. 2.7). 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 (fig. 2.7). 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.

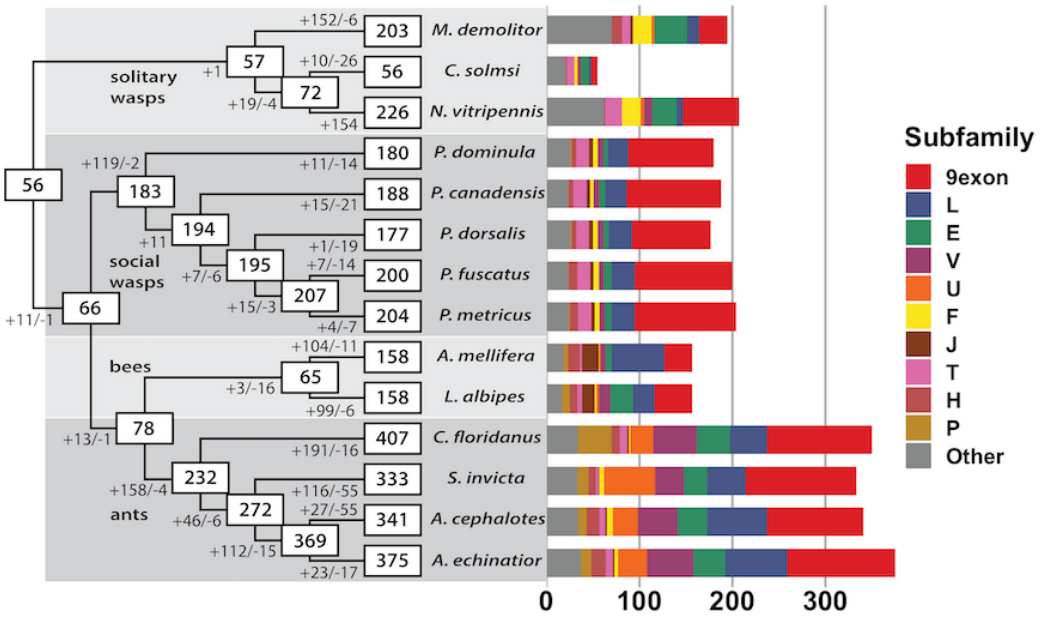

Figure 2.7. Cladogram of Hymenoptera species showing estimated number of OR gene gain and loss events along branches and estimated size of ancestral and extant species OR repertoires in boxes. To the right is a bar chart showing numbers of ORs broken down by subfamily. Non-Polistine OR data are from 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.

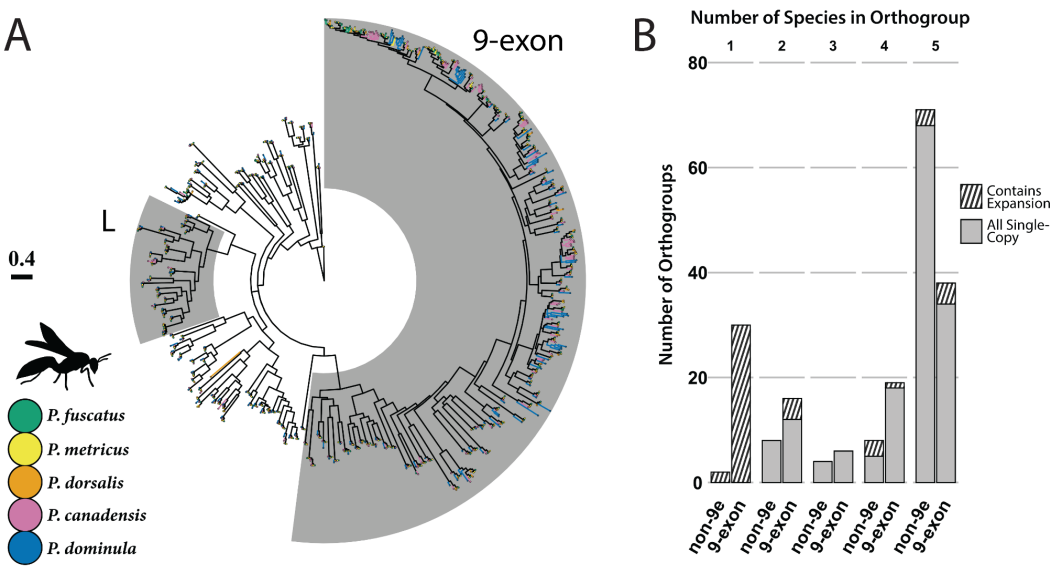

Figure 2.8. (A) Maximum likelihood OR protein tree with branches colored by species (Green: *Polistes fuscatus*; Yellow: *P. metricus*; Orange: *P. dorsalis*; Magenta: *P. canadensis*; Blue: *P. dominula*). The L and 9-exon subfamilies are highlighted. Scale bar represents 0.4 mean substitutions per 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).

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. 2.8A). 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) (table 2.3). The remaining orthologous groups contain an expansion in one or more species (fig. 2.8B). 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. 2.7). 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 2.3). 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. 2.8B). 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 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. 2.9A). 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 nonallelic 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. 2.9B). 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. 2.9B). 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. 2.9B). 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.

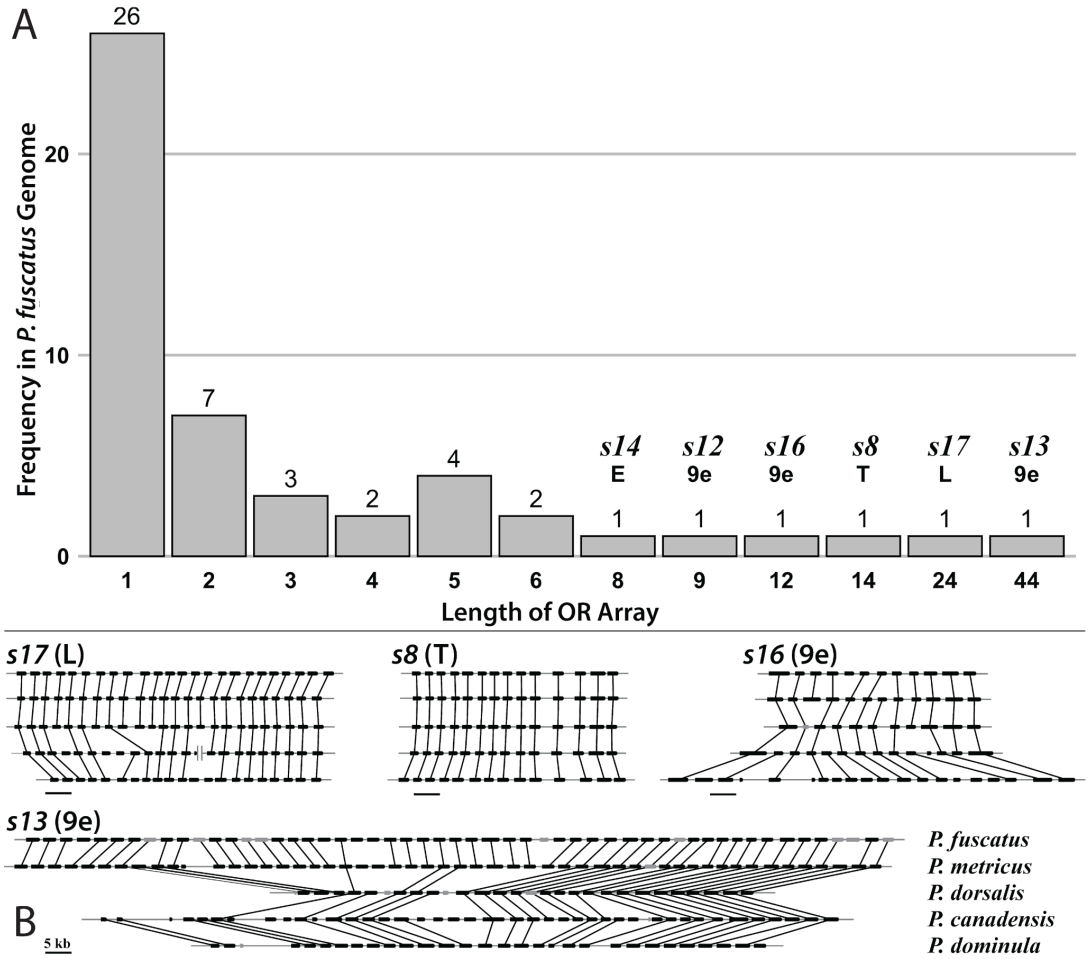

Figure 2.9. (A) Frequency of OR gene singletons and tandem arrays in the *Polistes fuscatus* genome. 62% of ORs in *P. fuscatus* occur in tandem arrays of six or more genes. The longest tandem array is a 44 gene cluster on scaffold 13 (s13) containing 9-exon subfamily ORs. The first row of x-axis labels 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.

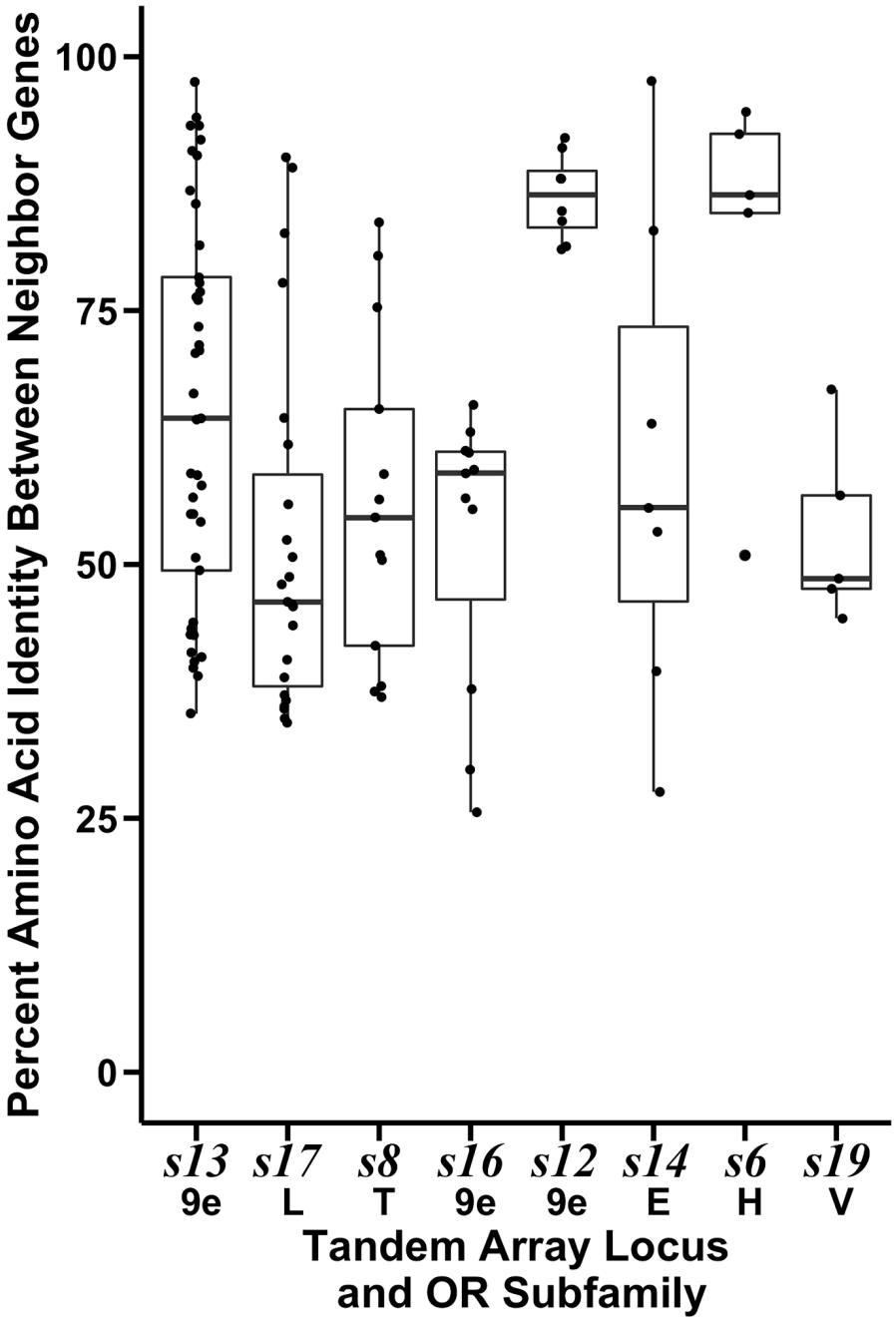

Figure 2.10. Percent amino acid identity between neighboring genes at eight loci containing the longest OR gene tandem arrays in the *P. fuscatus* genome. Arrays are ordered by length in gene number, from longest (44 9-exon subfamily ORs in the s13 tandem array) to shortest (6 H subfamily ORs in the s6 tandem array and 6 V subfamily ORs in the s19 tandem array).

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. 2.10). 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 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); three branches in the L subfamily (1.28% of 234 L subfamily branches); one branch in the E subfamily (1.67% of 60 E subfamily branches); and one branch in the H subfamily (1.54% of 65 H subfamily branches). 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 dN and dS for pairwise alignments of 151 single copy orthologs between *P. fuscatus* and *P. dorsalis* (fig. 2.11) using model yn00 of PAML (Yang 2007). Values of dN are significantly higher in 9-exon (mean dN = 0.015) compared to other OR ortholog pairs (mean dN = 0.006) (Welch Two Sample t-test; P-value = 5.317e−07). Values of dS are not significantly elevated among 9-exon ortholog pairs compared to other OR subfamilies (mean dS = 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) (fig. 2.12). 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.

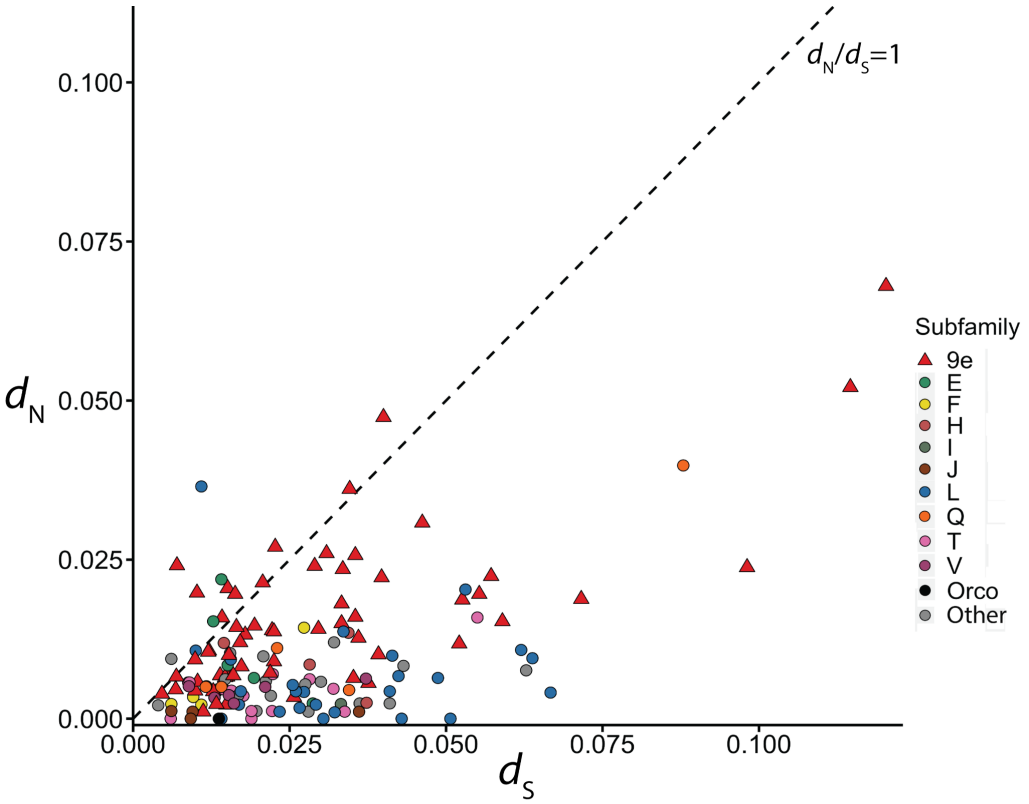

Figure 2.11. The values of dS (x-axis) and dN (y-axis) from pairwise alignments of *Polistes* fuscatus and *P. dorsalis* 1:1 orthologs. Values of dN are elevated in the 9-exon OR subfamily (data points represented by red triangles) relative to other OR subfamilies (data points represented by circles). The diagonal line represents a line of equality with slope of 1.

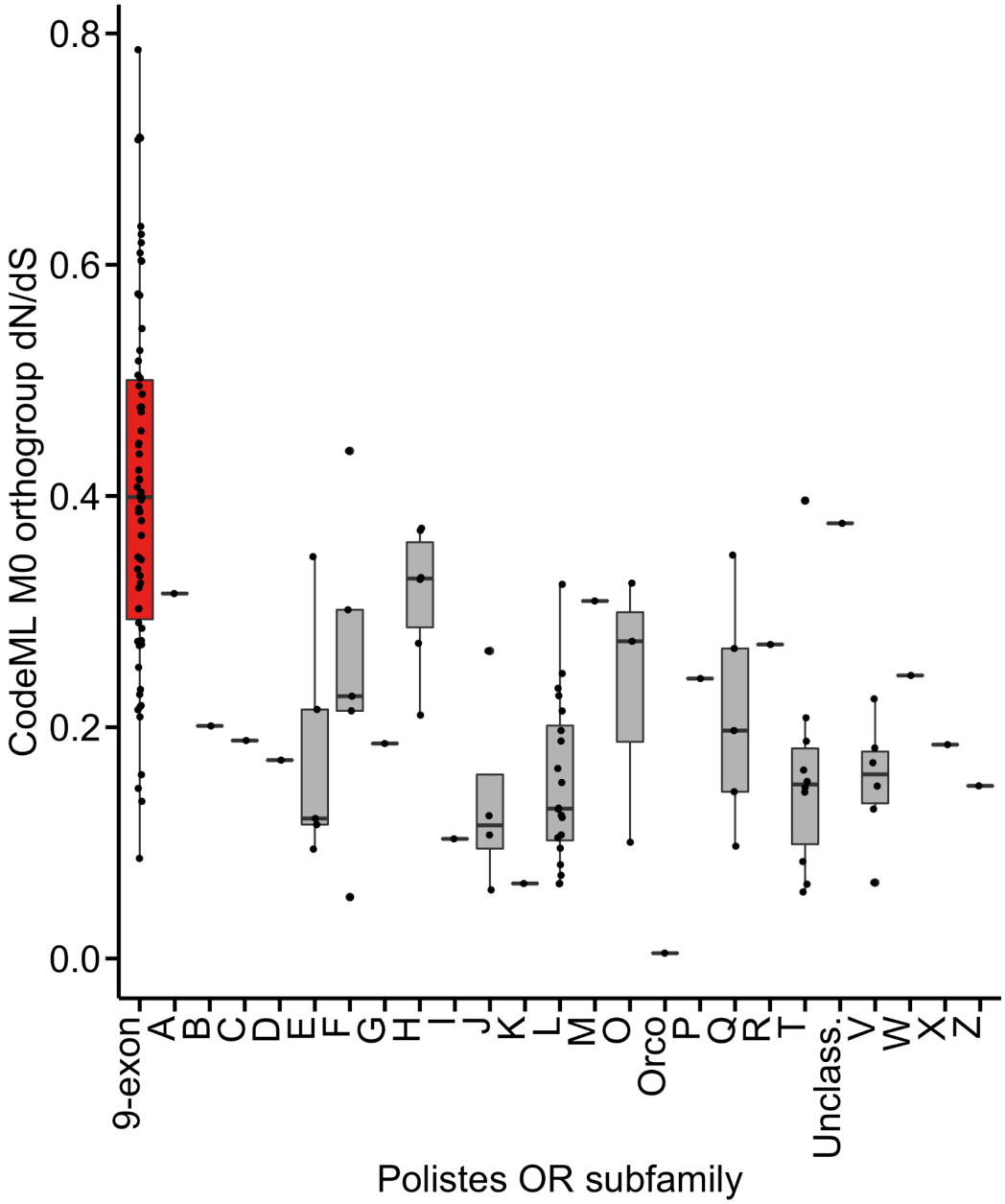

Figure 2.12. Box and whisker plot of omega values of 145 *Polistes* orthogroups as estimated by model M0 of CodeML. Values of omega are elevated in the 9-exon OR subfamily (red) relative to other OR subfamilies.

Discussion

Expansion of 9-Exon ORs 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 TB 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 TB 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. 2.9B) 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. 2.7). 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.

Acknowledgments

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]. 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.

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.

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

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.

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

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.

Fletcher W, Yang Z. 2010. The effect of insertions, deletions, and alignment errors on the branch- site test of positive selection. Mol Biol Evol. 27:2257-2267.

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.

Gao F, Chen C, Arab DA, Du Z, He Y, Ho SYW. 2019. EasyCodeML: a visual tool for analysis of selection using CodeML. Ecol Evol. 9:3891-3898.

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. 8:1494-1512.

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.

Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics. 30:1236-1240.

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 *Apis florea* 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.

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

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.

Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 47:W256-W259.

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.

Lopez-Osorio F, Pickett KM, Carpenter JM, Ballif BA, Agnarsson I. 2017. Phylogenomic analysis of yellowjackets and hornets (Hymenoptera: Vespidae, Vespinae). Mol Phylogenet Evol. 107:10- 15.

Loytynoja A, Goldman N. 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. Science. 320:1632-1635.

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.

Oxley PR, Ji L, Fetter-Pruneda I, Mckenzie SK, Li C, Hu H, Zhang G, Kronauer DJC. 2014. The genome of the clonal raider ant Cerapachys biroi. Curr Biol. 24:451-458.

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.

Pearson WR. 2013. Selecting the right similarity-scoring matrix. Curr Protoc Bioinformatics. 43:3.5.1-3.5.9.

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.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/.

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.

Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol. 3:217-223.

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. 2006. Phylogenetic models of rate heterogeneity: a high performance computing perspective. Proceedings 20th IEEE International Parallel and Distributed Processing Symposium, Rhodes Island, doi: 10.1109/IPDPS.2006.1639535.

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.

Yu G, Smith D, Zhu H, Guan Y, Lam TT-Y. 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8:28-36.

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.

CHAPTER 3

DETECTION OF SEX-BIASED GENE EXPRESSION IN PAPER WASP ANTENNAE

Andrew W. Legan, Michael J. Sheehan

Abstract

In insects, sex-dimorphic antenna morphology is commonplace, reflecting sex-correlated selective pressures shaping olfactory organ shape. Likewise, sex-specific antennal gene expression gives rise to sex-correlated behavioral responses to chemical cues and signals. In haplodiploid insects in the order Hymenoptera, females exhibit extreme behavioral complexity in the social ants, bees, and wasps. Males demonstrate comparatively limited behavioral complexity in the context of the colony. We examined sex-specific antennal gene expression in a facultatively polygynous social wasp, the northern paper wasp *Polistes fuscatus*, by sampling reproductive females and males. The most differentially expressed genes between the sexes were cytochrome P450 genes, which may play a role in xenobiotic or odorant degradation. We examined expression of the ~200 putatively functional odorant receptor genes in the *P. fuscatus* genome, half of which belong to an expanded 9-exon subfamily. While most 9-exon odorant receptors trend toward gyne-biased, antennal expression of 40% of 9-exon receptors are equally expressed in males and female or are male-biased, departing strongly from patterns observed in ants. The most male-enriched odorant receptor is a 9-exon receptor, *PfusOr94*. The conserved Hymenopteran tandem array of L subfamily odorant receptors encodes genes exhibiting generally male-biased antennal expression, although the most significantly female-biased odorant receptor is an L subfamily

receptor, *PfusOr185*. Patterns of sex-specific odorant receptor gene expression pinpoint candidate receptors involved in sex-specific functions, such as mate seeking.

Introduction

Selective pressures can differ between sexes, leading to sex-specific fitness peaks in sensory abilities (Lande 1980; Parsch and Ellegren 2013). Sex-dimorphic morphology in sensory systems is widespread among insects (Rospars and Chambille 1989; Nakanishi et al. 2009; Zhou et al. 2012; Zhao et al. 2016; Scherberich et al. 2017). Larger eyes in male carpenter bees accompany their visual mate search behavior (Somanathan et al. 2017). Drone (male) honeybee antennae are larger than female antennae, aiding in olfactory localization of queens during mating flights of competing drones (Koeniger et al. 2005; Wanner et al. 2007b). Sexual selection also shapes molecular evolution of sensory receptor genes, such as chemoreceptors.

Insect multigene chemoreceptor families encode sensory receptor proteins diverse in copy number, sequence, and pattern of gene expression (Robertson et al. 2003; Benton 2015; McKenzie and Kronauer 2018). Odorant receptors (ORs) are heteromeric cation channels in the membranes of olfactory receptor neurons in sensory organs, such as the antennae (Suh et al. 2014; Zhao and McBride 2020). In insects, sex-specific expression of odorant receptors contributes to sex-specific behavioral responses to odorants. For example, in the silkworm *Bombyx mori*, host plant-seeking by females is mediated by the female biased odorant receptors *BmOr19* and *BmOr30*, and mate seeking by males is mediated by the male biased *BmOr1* and *BmOr3* (Wanner et al. 2007a; Nakagawa et al. 2005). In part because of interest in female-specific

complex behavior of social insects, our understanding of sexual selection in the social Hymenoptera is limited.

In Hymenoptera, sex determination results from haplodiploidy, where males generally develop from an unfertilized egg and are haploid while females develop from a fertilized egg and are diploid. Social Hymenoptera colonies function by the collective behavior of females, and males contribute little to the social coordination of the colony. Instead, males typically eclose late in the colony and are short-lived, focusing exclusively on reproduction. For example, in the honeybee *Apis mellifera*, the hallmarks of male behavior are its "simplicity" and "limited diversity" (Ohtani 1974). How have divergent selective regimens between sexes shaped social insect olfactory systems at the molecular level? Insect odorant receptor genes contribute to variation in olfactory behavior between individuals and provide a good set of candidates for sex-specific sensory gene expression (Rihani and Sachse 2022).

Large complements of odorant receptor genes (ORs) have evolved among social Hymenoptera, giving rise to perception of complex chemical cues and signals (Smith CR, Smith CD et al. 2011; Smith CD, Zimin et al. 2011; Zhou et al. 2012; Engsontia et al. 2015; Zhou et al. 2015; Karpe et al. 2016; McKenzie et al. 2016; Legan et al. 2021). Hymenopteran odorant receptors are divided into subfamilies based on their sequence similarity, and these subfamilies have been assigned putative functions based on their electrophysiological responses to chemical stimuli in the honeybee and some ant species. For example, an odorant receptor in *Apis mellifera* (*AmOr11*) responds to components of the queen mandibular pheromone (Wanner et al. 2007b). *AmOr11* is in the L subfamily of Hymenopteran odorant receptors. In *A. mellifera* and *Apis florea*, L

subfamily OR antennal expression is frequently biased towards male drones compared to female workers (Wanner et al. 2007b; Karpe et al. 2016). L and 9-exon OR subfamily receptors, as well as ORs in at least six other subfamilies, respond to cuticular hydrocarbons in functional characterization experiments (Pask et al. 2017; Slone et al. 2017).

In social Hymenoptera, cuticular hydrocarbons (CHCs) are frequently signals or cues that inform social behavior. Behavioral observations of limited social behavior in male social Hymenoptera warrant the prediction that males will express a limited range of odorant receptors attuned to social chemicals like CHCs. Indeed, antennal gene expression of 9-exon ORs tends to be biased towards female workers relative to males in ants and bees (Zhou et al. 2012; McKenzie et al. 2016; Karpe et al. 2016). However, L subfamily receptors also respond to CHCs in functional characterization studies (Slone et al. 2017). These L subfamily ORs tend to show male-biased antennal gene expression in ants and bees (Wanner et al. 2007b; Zhou et al. 2012; McKenzie et al. 2016; Karpe et al. 2016). In *Polistes*, the 9-exon OR subfamily has independently expanded to around 100 genes (Legan et al. 2021). The L subfamily is also a considerable size at 24 genes, all within an ancient tandem array ancestral to ants, bees, and wasps (McKenzie and Kronauer 2018; Legan et al. 2021). Sex-specific odorant receptor expression, or antennal gene expression more generally, has not been previously investigated in social wasps.

Features of the *Polistes* mating system and male behavior make paper wasps an intriguing clade for the study of sexual selection (Beani et al. 2014). Reproductive females (gynes) emerge in the late summer and early fall, mate, and overwinter before

founding nests in the spring (West-Eberhard 1969; Reeve 1991). These gynes do not forage and are typically found on the nest (West-Eberhard 1969). Males emerge concurrent with gynes (West-Eberhard 1969). In multiple *Polistes* species, males aggregate in leks, where they wait for females to visit to mate (Post and Jeanne 1983a, 1983b, 1984; video S1). Males employ multiple mate search strategies, including aggregation at leks, territory defense and patrolling, and ranging (Beani and Turillazzi 1988; Beani et al. 2019). In *P. dominulus*, males employ visual signals of competitive ability in the form of spheroid color patterns on the abdomen (Izzo and Tibbetts 2012; de Souza et al. 2017b). Male aggregations can become large in some contexts, and territorial defense can involve extreme aggression. In the context of the nest, male *P. fuscatus* and *P. metricus* will sometimes feed larvae, a behavior typically assumed to be expressed only in workers and queens, however they do not forage (Hunt and Noonan 1979; see video).

There is evidence that *Polistes* males and gynes depend on their sense of smell during the mating season. In *P. exclamans*, both males and reproductive females (gynes) are attracted to chemicals of the opposite sex (Reed and Landolt 1990; Elmquist et al. 2018). Male *P. dominula* likely discriminate between workers and gynes, and between parasitized and unparasitized gynes, based on chemical cues (Cappa et al. 2013). These results suggest that in *Polistes*, both males and females rely on olfaction to find and assess mates. However, the mate attractant chemicals may differ between sexes, as there is moderate sexual dimorphism in cuticular hydrocarbons between males and females (Espelie and Hermann 1990; Layton et al. 1994; Beani et al. 2019). There is also some sexual dimorphism in the shape and composition of

Polistes antennae. Male antennae house several glands absent in female antennae in *P. dominula* (Romani et al. 2005). Male *Polistes* have an additional antennal segment compared to females and are typically equal or smaller in body size compared to females (Miller and Sheehan 2021). Distal segments of *Polistes* male antennae curl to form a hook at the antenna terminus (Pekkarinen and Gustafsson 1999). *Polistes* male courtship behavior involves tapping females with their antennae and curling their antennae around the female's antennae during copulation (West-Eberhard 1969; Romani et al. 2005).

Given the sex-specificity of antennal morphology, antennal gland composition, CHC profiles, mate search behavior, and courtship antennation, we hypothesized that selection has favored sex-specific olfactory perception in *Polistes* and predicted sex-specific gene expression in *Polistes* antennae. To test this prediction, we sequenced the antennal RNA of the northern paper wasp *Polistes fuscatus*. Rather than comparing males to non-reproductive workers, we compared males to reproductive females (gynes), eliminating the confound of reproductive status present in comparisons between males and workers.

Methods

Six *P. fuscatus* males and six reproductive females (gynes) were collected in the fall of 2017 in Ithaca, NY and were housed in deli cups in the lab, with water vials and a sugar cube in the deli cups. All wasps were provided an opportunity to mate. Each wasp was collected from a different nest, except one female and one male that were collected from the same nest. Antennae were resected from heads and flash frozen in liquid nitrogen, then stored at -80°C until RNA isolation using the Qiagen RNeasy Mini

kit according to the manufacturer's recommendations. Each RNA sample represents the total RNA isolated from two pooled antennae of each individual. Libraries were prepared in house using the NEBNext Ultra II RNA Library Prep Kit for Illumina (NEB #E7770L) with the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB #E7490) and NEBNext Multiplex Oligos for Illumina (NEB #E7600S) and sequenced on two lanes of an Ilumina HiSeq (Illumina, San Diego, CA, USA). Raw reads were trimmed using Trimmomatic v.0.39 (Bolger et al. 2014) as described in Legan et al. (2021). Trimmed reads (average length 132 bp) were then mapped to the *P. fuscatus* genome (Miller et al. 2020) using STAR v.2.7.9 (Dobin et al. 2013).

Software bedtools "multicov" function was used to generate read counts per gene, treating all reads as single end reads and counting reads that mapped to exon regions of genes (Quinlan and Hall 2010). Sex-dimorphic expression was quantified in R version 3.5 using the edgeR package (Robinson et al. 2010; R Core Team 2018). Because male antennal RNA expression sequenced in sample "M50" was an extreme outlier in relation to other male samples, and in relation to female samples (fig. 3.1), this sample was excluded from analysis. The edgeR "glmLRT" function was used to test for differential expression of all annotated genes (11,805 genes) using likelihood ratio tests (Chen et al. 2008). P-values reported represent Benjamini and Hochberg adjusted P-values (Benjamini and Hochberg 1995), output of the edgeR function "topTags". Gene ontology (GO) analysis was done using the topGO package in R using the "classic" method and Fisher's exact test (Alexa and Rahnenfuhrer 2021). Weighted gene co-expression network analysis was performed with R package WGCNA version 1.71 (Langfelder and Horvath 2008).

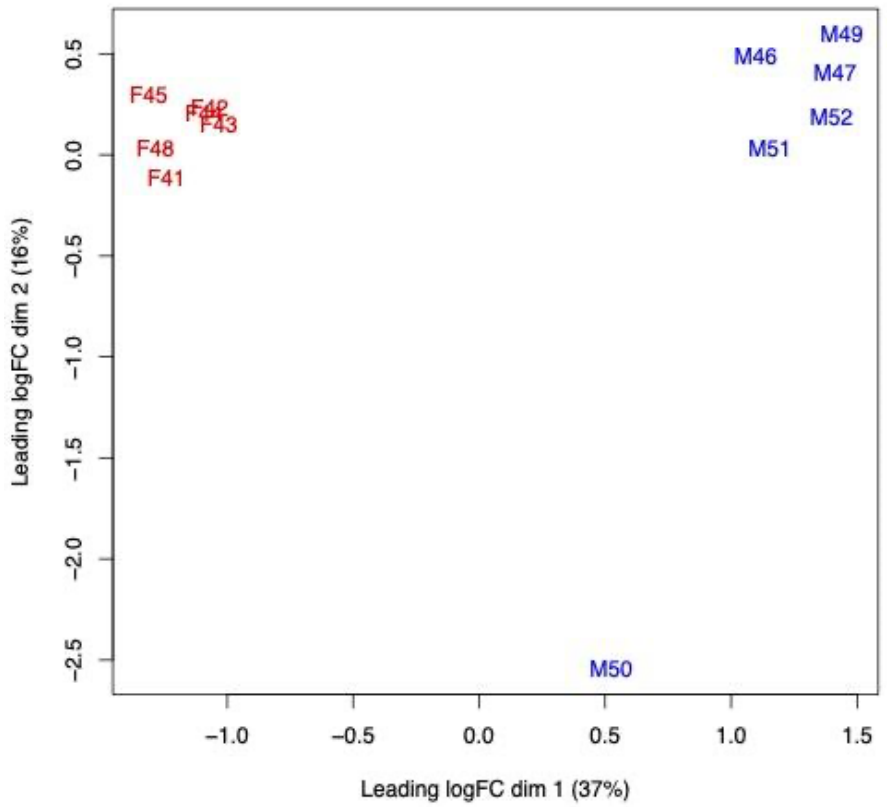

Figure 3.1. Multidimensional scaling plot of differences in antennal gene expression between samples.

Odorant receptor gene and pseudogene models encoding proteins with lengths less than 350 amino acids were excluded from phylogenetic reconstruction. Protein sequences were aligned using MAFFT (Katoh and Standley 2013) version 7.453 with options "--genafpair" and "--maxiterate 1000" (strategy E-INS-i). The alignment was trimmed using trimAl with option "-gappyout" to remove gap rich columns, generating a phylip alignment of 203 sequences with length 334 amino acids (Capella-Gutiérrez et al. 2009). A phylogenetic tree of odorant receptor protein sequences was constructed with RAxML version 8.2.12 using the maximum likelihood optimality criterion with 100 bootstrap inferences and the odorant receptor coreceptor as outgroup (Jones et al. 1992; Stamatakis 2014). The tree was visualized in R using the ggtree package and expression measures mapped with the ggtree "gheatmap" function (Yu et al. 2017).

Sex-specific odorant receptor expression data for other Hymenoptera species were accessed in online supplementary materials. Data from the ants *Camponotus floridanus*, *Harpegnathos saltator*, and *Ooceraea biroi* were accessed from supplementary table S7 in Zhou et al. (2015). Data from the honeybee *Apis florea* were accessed from supplementary table Suppl-2 in Karpe et al. (2016). Phylogenetic reconstruction of Hymenopteran ORs in 9-exon and L subfamilies was carried out using the same parameters for alignment, trimming, and maximum likelihood tree creation as described above, with no outgroups specified. Amino acid sequences for the honeybee *Apis florea* were accessed from supplementary file Suppl-1 in Karpe et al. (2016) and protein sequence data for two ants, *C. floridanus* and *H. saltator*, were accessed from Dataset S1 in Zhou et al. (2012). Only those odorant receptors with predicted protein sequences greater than 350 amino acids in length were used.

Results

Sex-specific antennal expression of cypP450s, muscle-related, and other genes in *P. fuscatus*:

P. fuscatus adults were distinctly separated by sex along the first MDS axis which explained 44% of the variation in antennal gene expression between samples (fig. 3.2). A subset of 1,128 genes out of 10,688 expressed genes (10.6% of expressed genes) were significantly differentially expressed at Benjamini and Hochberg adjusted significance level of $P < 0.01$ (fig. 3.3; fig. 3.4). In both males and females, the top differentially expressed gene was a cytochrome P450. Twenty-six out of the top fifty male-biased genes (52%) and fifteen out of the top fifty female-biased genes (30%)

were predicted cytochrome P450 genes. Three of the top fifty male-biased genes were fatty acid desaturases.

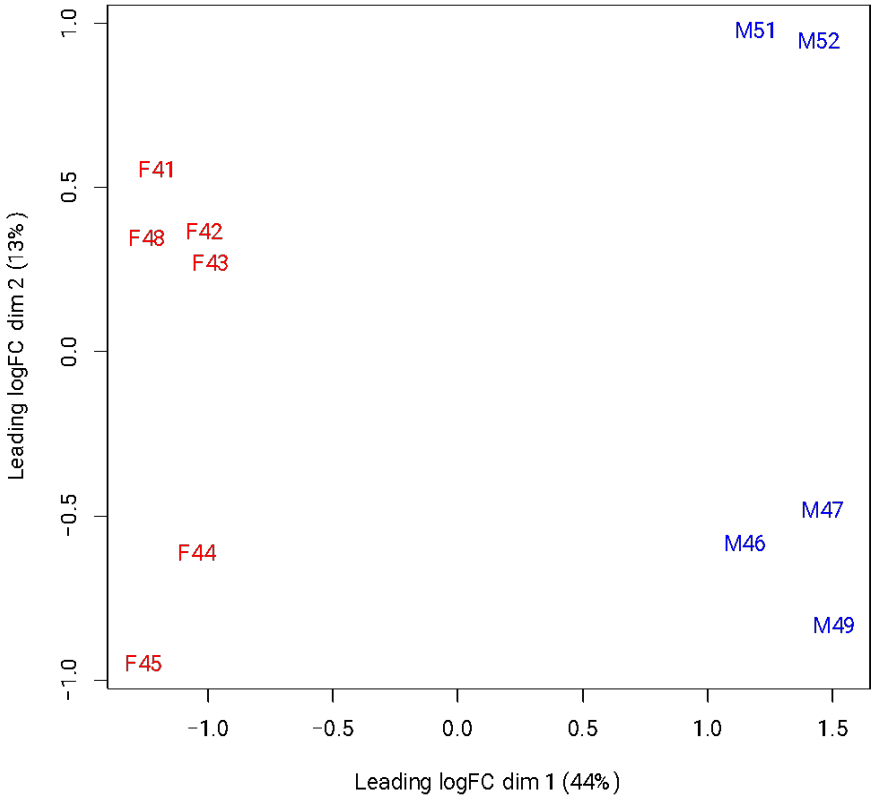

Figure 3.2. Multidimensional scaling plot of differences in antennal gene expression between samples after removal of aberrant male sample M50.

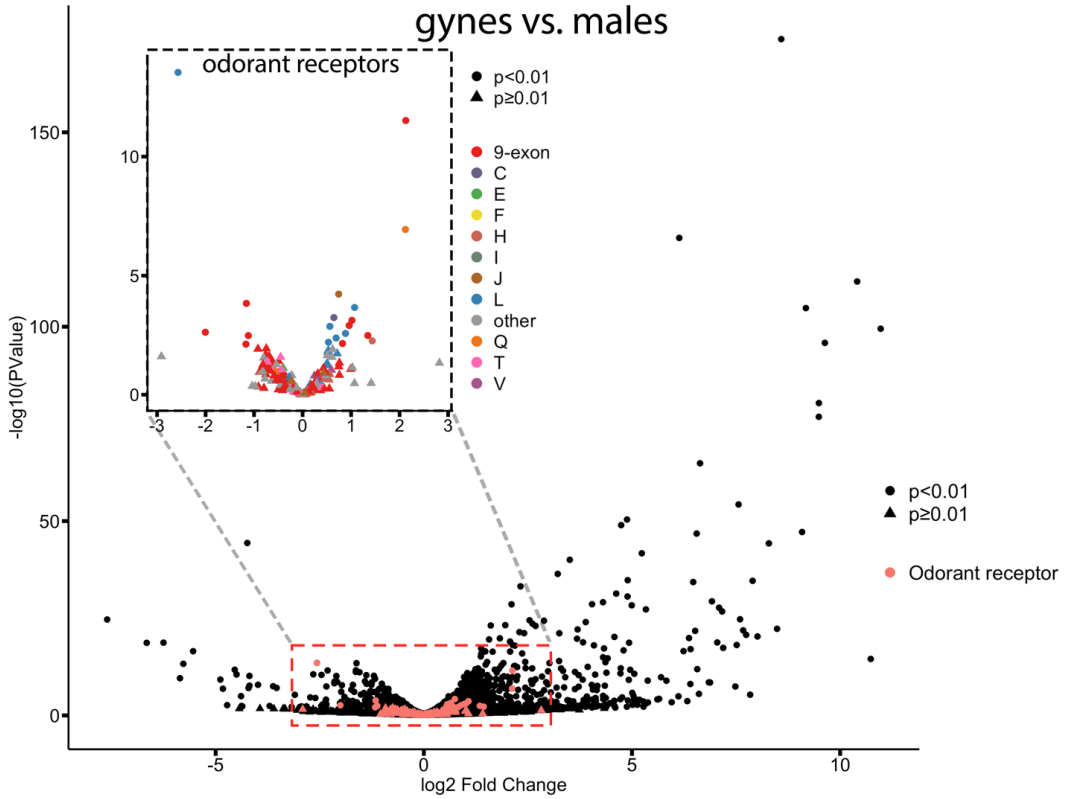

Figure 3.3. Volcano plot showing sex-dimorphic antennal gene expression analysis of all annotated genes in the *P. fuscatus* genome. Inset shows only the odorant receptor genes, color-coded by odorant receptor subfamily. ORs with positive fold change values are enriched in male antennae, and those with negative fold change values are enriched in female (gyne) antennae. Genes represented by circles were differentially expressed at a Benjamini and Hochberg adjusted significance threshold of $P < 0.01$.

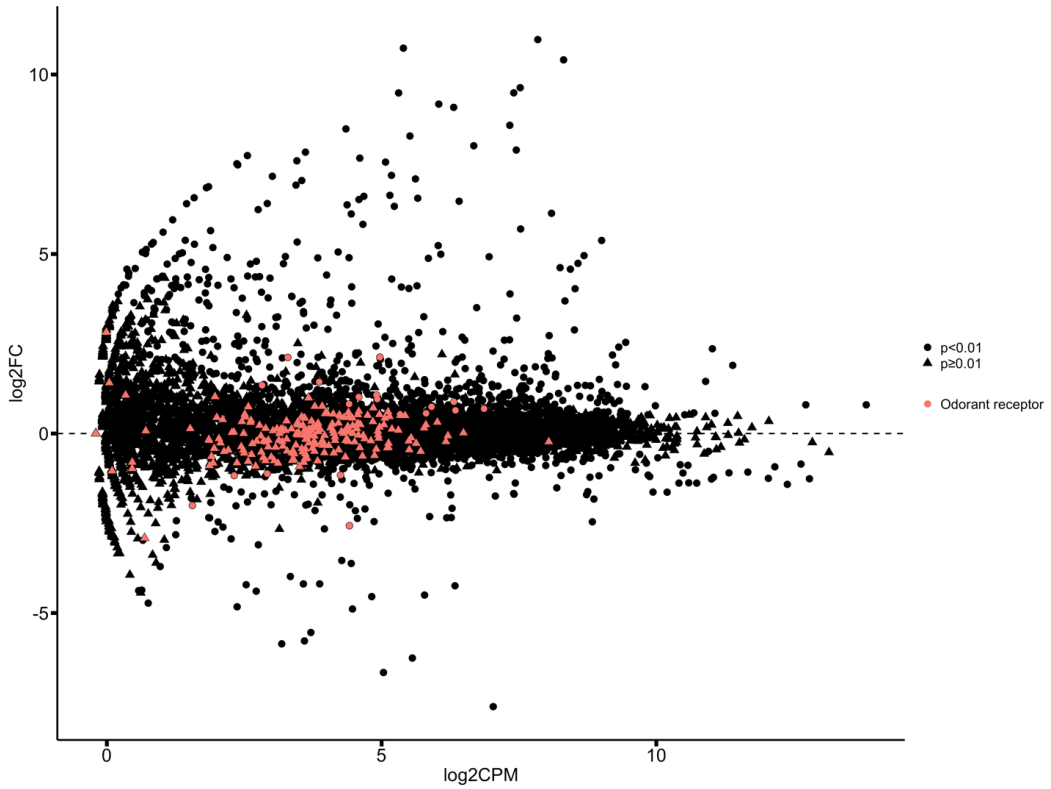

Figure 3.4. MA plot with *Polistes fuscatus* tuning odorant receptors overlaid. The x-axis represents the base mean expression of each gene, measured as log2(counts per million). The y-axis represents fold change in gene expression between males and gynes, measured as log2(FC), where positive values indicate male-biased gene expression and negative values indicate gyne-biased gene expression.

Gene ontology (GO) term analysis identified 17 terms (of 3031 terms annotated) significantly enriched for 317 sex-biased genes at $P < 0.01$, including six terms (89 genes) associated with muscle tissue (table 3.1). Long-chain fatty acid metabolic process genes were enriched in the GO analysis. Examination of these six genes revealed three elongation of very long chain fatty acids proteins, two 4-coumarate--CoA ligase 1-like genes, and one long-chain-fatty-acid--CoA ligase ACSBG2. Organophosphate catabolic process was also an enriched GO category (10 significant genes), which may relate to the breakdown of human-made organophosphate

components of agricultural insecticides, and 7 of these genes displayed male-biased expression.

Table 3.1. Seventeen gene ontology term categories enriched for sex-specific gene expression.

Term	Annotated	Significant	Classic Fisher
heart contraction	48	8	3.20E-05
gliogenesis	87	13	0.00031
sarcomere organization	61	19	0.0005
myofibril assembly	89	30	0.00054
muscle system process	44	17	0.00208
organophosphate catabolic process	35	10	0.00253
regulation of establishment of planar polarity	23	8	0.0034
long-chain fatty acid metabolic process	12	6	0.00341
inorganic anion transmembrane transport	23	8	0.00468
muscle cell cellular homeostasis	29	10	0.00479
compartment pattern specification	13	6	0.00558
larval visceral muscle development	10	5	0.00764
glucose homeostasis	45	13	0.00768
plasma membrane bounded cell projection organization	912	146	0.00776
negative regulation of stem cell differentiation	18	7	0.00853
aminoglycan catabolic process	20	6	0.00978
cGMP metabolic process	17	5	0.00981

Patterns of gene expression in the *Polistes fuscatus* odorant receptor gene family:

The bulk of the odorant receptor genes display average expression levels in comparison with the rest of the annotated genes in the *P. fuscatus* genome (fig. 3.4). The odorant receptor coreceptor displayed the highest expression (log(CPM) = 9.7) of any odorant receptor, consistent with its function as a requisite receptor subunit expressed ubiquitously in olfactory receptor neurons. The top expressed tuning odorant receptor is in the J subfamily, *PfusOr95*. About half of *Polistes* odorant receptors belong to an expanded clade of 9-exon subfamily ORs (Legan et al. 2021). In *P.*

fuscatus, many 9-exon ORs exhibit lower expression relative to the rest of the OR subfamilies (fig. 3.5). However, some 9-exon ORs at the base of the expanded subfamily clade are highly expressed in both sexes (fig. 3.5). L subfamily ORs are overall expressed more in both gyne and male antennae relative to other OR subfamilies (fig. 3.5).

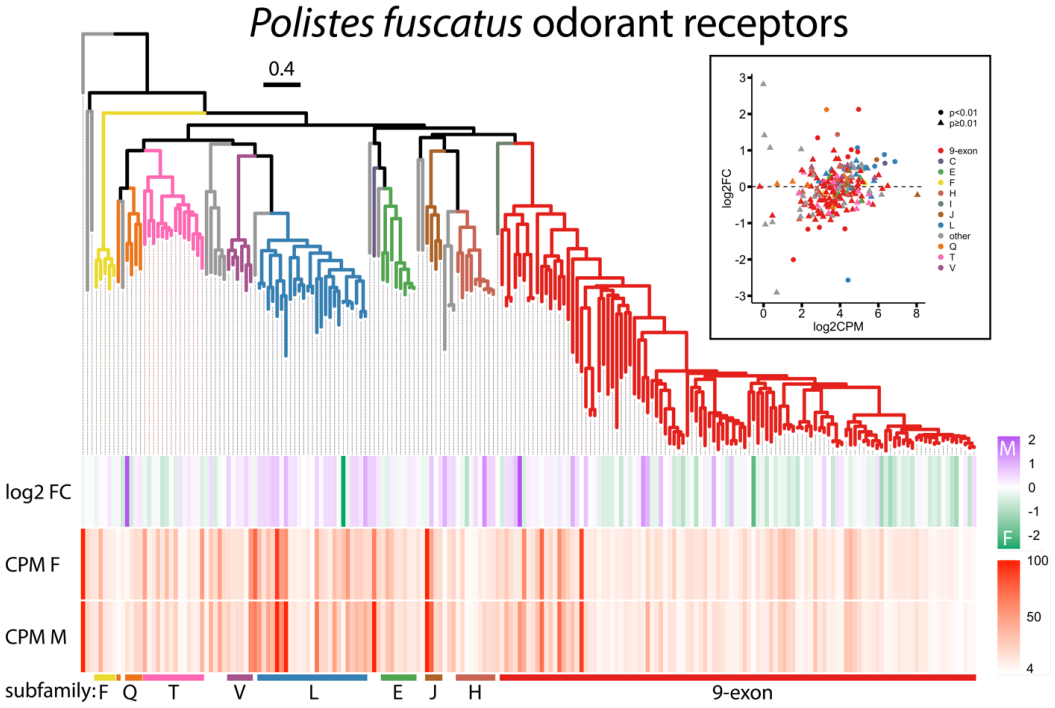

Figure 3.5. *Polistes fuscatus* odorant receptor gene tree, based on amino acid alignment, with gene expression data for individual genes mapped to the tree's branch tips. Scale bar at the top represents estimated divergence in amino acid substitutions per site. Tree branches are color coded by select odorant receptor subfamily clades as in fig. 3.3 (inset). CPM represents the averages of counts per million per gene across six females (CPM F) and five males (CPM M). Upper right panel shows MA plot of *P. fuscatus* ORs.

Sex-biased antennal expression of odorant receptor genes in *Polistes fuscatus*:

Nineteen odorant receptors (ORs) were differentially expressed between males and gynes at Benjamini and Hochberg adjusted significance level of $P < 0.01$ (fig 3.3 inset). In gyne antennae, five ORs were enriched compared to males. These include four 9-exon ORs (*PfusOr84*, *PfusOr128*, *PfusOr143*, *PfusOr155*), and one OR in

subfamily L (*PfusOr185*). Fourteen ORs were significantly enriched in male antennae compared to gynes. These include five ORs in the L subfamily (*PfusOr179*, *PfusOr180*, *PfusOr186*, *PfusOr192*, *PfusOr195*), five ORs in the 9-exon subfamily (*PfusOr19*, *PfusOR94*, *PfusOr107*, *PfusOr221*, *PfusOr222*), and one OR each in subfamilies C (*PfusOr204*), H (*PfusOr48*), J (*PfusOr99*), and Q (*PfusOr98*). Weighted gene co-expression network analysis (WGCNA) identified 39 modules of co-expressed genes (fig. 3.6). Odorant receptor genes are spread across 35 WGCNA modules (table S3.1).

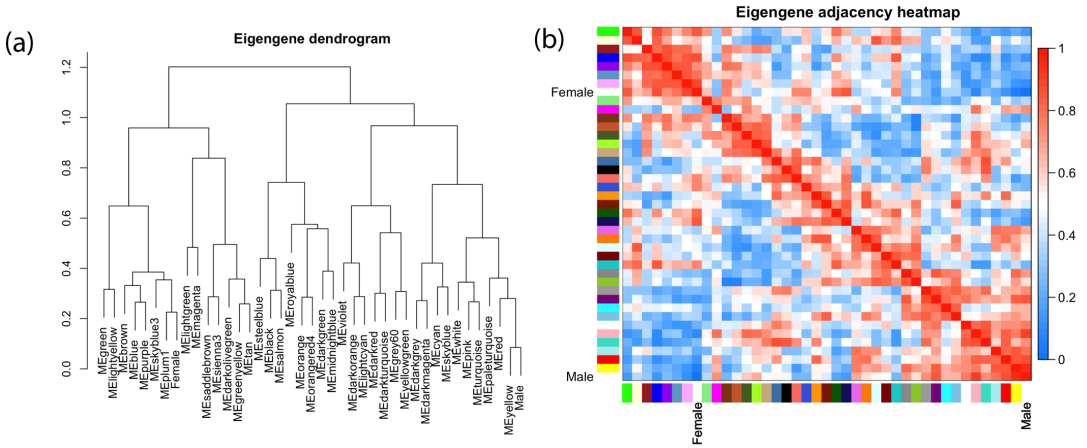

Figure 3.6. Results of WGCNA analysis. (a) Hierarchical clustering of eigengene modules and sample traits male and female. (b) Heatmap of eigengene adjacencies. Red corresponds to high correlation between modules and blue corresponds to negative correlation between modules.

Sex-biased OR gene expression across social Hymenoptera:

Patterns of sex-specific gene expression of the 9-exon and L subfamilies follow trends observed in other social Hymenoptera species, with 9-exon ORs frequently female biased and L subfamily ORs frequently male biased (fig. 3.7). However, the *P. fuscatus* 9-exon subfamily is not gyne biased to the degree that this subfamily is worker-biased in a honeybee and three ant species (fig. 3.7A). Moreover, a larger proportion of the *P. fuscatus* L subfamily is male biased than that of the other four social Hymenopterans examined (fig. 3.7B).

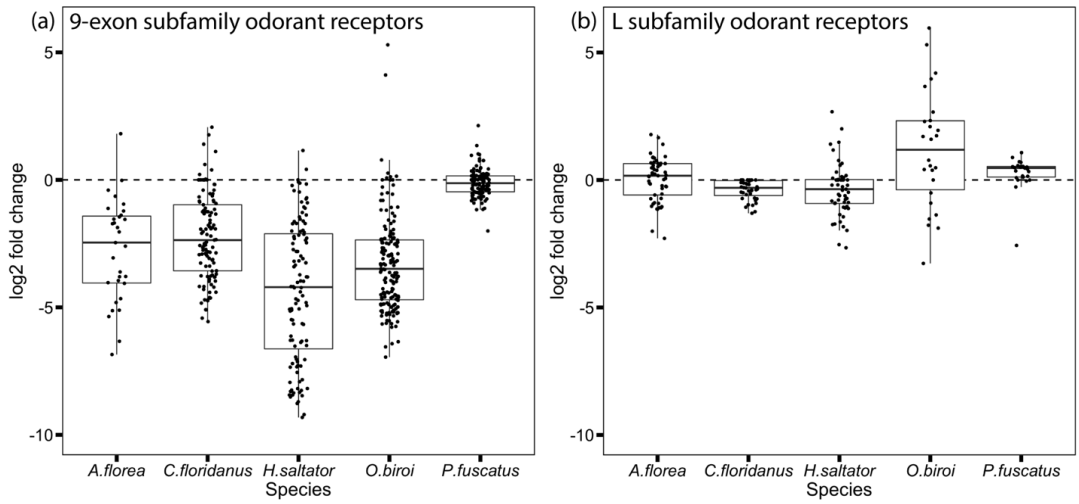

Figure 3.7. Sex-dimorphic antennal gene expression (log2 fold change in gene expression) for subfamily 9-exon (a) and subfamily L (b) odorant receptors in the honeybee *Apis florea* (data from Karpe et al. 2016), three ant species (*Camponotus floridanus*, *Harpegnathos saltator*, and *Ooceraea biroi*; data from Zhou et al. 2015), and the paper wasp *Polistes fuscatus* (this study). ORs with positive fold change values are enriched in male antennae, and those with negative fold change values are enriched in female antennae. The dashed horizontal line denotes zero-fold change.

Lineage-specific expansions of 9-exon ORs in *A. florea*, *C. floridanus*, and *H. saltator* tend to display worker biased antennal expression (fig. 3.8A). The relatively few male-biased 9-exon ORs in *A. florea*, *C. floridanus*, and *H. saltator* are clustered in clades exhibiting more cross-species orthology near the base of the gene tree (fig. 3.8A). Male-biased ORs are scattered throughout the gene tree of L subfamily odorant receptors (fig. 3.8B). *AfOr11*, ortholog of queen pheromone receptor *AmOr11*, is within a basal clade of honeybee L subfamily ORs and does not exhibit clear orthology with ant or wasp ORs (fig. 3.8B).

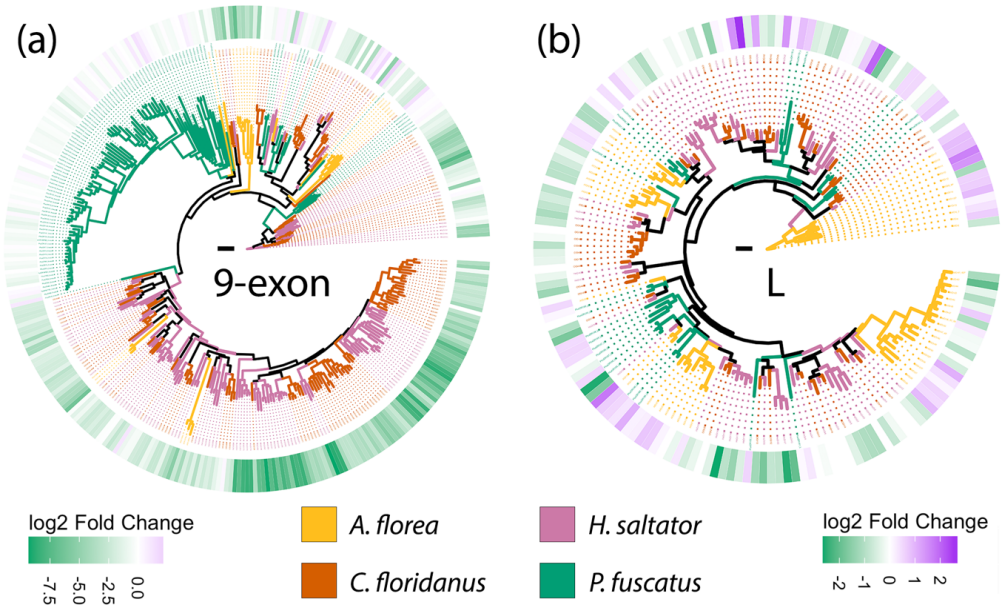

Figure 3.8. Phylogenetic reconstruction based on amino acid alignments of (a) 9-exon subfamily odorant receptors and (b) L subfamily odorant receptors from the honeybee *Apis florea* (data from Karpe et al. 2016), two ant species (*Camponotus floridanus*, *Harpegnathos saltator*; data from Zhou et al. 2012, 2015), and the paper wasp *Polistes fuscatus* (data from Legan et al. 2021, this study). Sex-biased gene expression data is mapped onto the trees as log2(fold change) where negative values (green) are gyne biased and positive values (purple) are male biased. Scale bars represent estimated divergence in amino acid substitutions per site (0.3).

Discussion

We identified patterns of differential gene expression in the antennae of reproductive female (gyne) and male social wasps *Polistes fuscatus* using bulk mRNA sequencing. Cytochrome P450 genes displayed stark contrasts in mRNA levels between sexes. We identified fourteen male-enriched and five gyne-enriched odorant receptor genes. These differentially expressed genes may contribute to sex-specific olfactory perception through enzymatic silencing of odorants by P450 enzymes and neuronal excitation after ligand-binding by odorant receptor channels.

Odorant degradation enzymes contribute to inactivating signals from lingering odorants and to lessen sensory noise resulting from overabundance of odorants in the

antennae (Ishida and Leal 2005; Durand et al. 2011; Leal 2013). Sex-specificity of odorant degradation enzymes has been observed in the cotton leafworm (Durand et al. 2011). A honeybee worker-biased cytochrome P450 gene was posited to play a role in odor degradation (Wanner et al. 2007b). Cytochrome P450 genes exhibited caste-specific expression in termites (Shigenobu et al. 2022). The frequent sex-specificity in cytochrome P450 gene expression in *Polistes* antennae may reflect sex-specific patterns of P450 enzymatic odorant degradation activity. However, P450 enzymes constitute a large and diverse gene family with multiple functions in degradation and synthesis, including xenobiotic degradation, modulation of juvenile hormone synthesis, cuticular hydrocarbon synthesis, and other functions (Sutherland et al. 1998; Feyereisen 2006; Qiu et al. 2012). Our findings raise the intriguing possibility of sex-specific odorant perception in paper wasps mediated by odorant degrading enzymes, though this idea will need to be tested in future work.

In social Hymenoptera, females are involved in more complex social interactions than males, while males usually have a more singular focus on mating. These differences in social behaviors appear to correlate with sex-differential expression of ORs in ants and bees. 9-exon subfamily ORs, and many ORs in other subfamilies, are generally enriched in female worker ant antennae compared to males (Wanner et al. 2007b; Zhou et al. 2015; McKenzie et al. 2016). Our observation departs from observed patterns, as we found five significantly male-biased 9-exon receptors compared to gynes. Still, the majority of *P. fuscatus* 9-exon ORs trend towards female-biased (Figure 2), and 9-exon ORs were generally lowly expressed in both sexes.

Our observation of generally male-enriched L subfamily ORs echoes results from social bees and ants. While solitary bees possess one third the number of L subfamily receptors as honeybees, and parasitoid wasp *N. vitripennis* genome encodes few L subfamily ORs, the L subfamily is frequently expanded in social Hymenoptera (Karpe et al. 2017). The honeybee L subfamily odorant receptor, *AmOr11*, is a drone-biased sex pheromone receptor, responding to the queen substance component 9-ODA (Wanner et al. 2007b). Several *H. saltator* L subfamily ORs responded to linear alkanes (Slone et al. 2017). In *P. fuscatus*, L subfamily ORs were some of the highest expressed ORs in both sexes, indicating important function for males and females alike.

Sex-differential gene expression of odorant receptors in social Hymenoptera has mainly been explored by comparing reproductive males to non-reproductive workers (Wanner et al. 2007b; Koch et al. 2013; Zhou et al. 2012; McKenzie et al. 2016). However, differentially expressed genes identified using this comparison vary between sex and reproductive status, and therefore may represent differences in reproductive status rather than differences in sex per se. Indeed, recent transcriptomic analysis of the termite *Reticulitermes speratus* found that head gene expression of reproductive males and reproductive females is more similar than between reproductive individuals and either males or females of soldier or worker classes (Shigenobu et al. 2022). A common tuning of the antennal olfactory system towards pheromones and chemical cues important for reproduction may reflect the relative lack of sex-specific expression of 9-exon ORs in *P. fuscatus*. As reproductive females move into different roles as foundresses and later queens, they may experience changes in odorant receptor abundances. Differences in brain morphology and brain gene expression correlate with

paper wasp caste differences and differences in social experience (O'Donnell et al. 2014; Jernigan et al. 2021; Uy et al. 2021). Comparisons of OR expression across life stages will be useful to explore the relationship between wasp developmental stage and olfaction.

Polistes males and gynes likely rely on olfaction to locate mates. There are differences in the cuticular hydrocarbons of different sexes in *Polistes*, so the chemicals used in mate choice may differ between sexes (Espelie and Hermann 1990; Layton et al. 1994; Beani et al. 2019; da Silva et al. 2021). The role of sexual selection in sex-dimorphism in CHC profiles is understudied in *Polistes*, but sex-specific CHCs provide good candidate ligands for the sex-specific ORs identified in this study (Berson et al. 2019). Male and female *P. exclamans* are attracted to hexane extracts of the opposite sex in a wind tunnel (Reed and Landolt 1990). *P. fuscatus* males are attracted to contents of the female venom gland of both *P. fuscatus* and *P. exclamans* and rely on cuticular chemicals of the female in mate compatibility recognition (Post and Jeanne 1984). Odorant receptors influencing mate choice may be mechanisms of behavioral reproductive barriers, potentially influencing speciation (Brand et al. 2015; Xu and Shaw 2019). Both male and gyne *P. fuscatus* discriminate between conspecifics and heterospecifics of the closely related *P. metricus*, likely mediated in part by chemical cues or pheromones (Miller et al. 2019).

Due to alternate female dispersal strategies, where some females disperse long distances while others remain in the general vicinity of their nest of origin (Bluher et al. 2020), males and gynes may benefit by discriminating between related and unrelated mates. Fitness in both sexes benefits from outbreeding by reducing the rate of diploid

male production, a costly phenomenon in haplodiploid social Hymenoptera (Boomsma et al. 2009). There is evidence both for (Ryan and Gamboa 1986) and against (Larch and Gamboa 1981; Post and Jeanne 1983b; de Souza et al. 2017a) the notion that *Polistes* males can recognize female nestmates, and this ability may vary between species. Ryan and Gamboa (1986) demonstrated that *P. fuscatus* "males and gynes from the same colony (i.e., nestmates) are significantly more tolerant of each other but copulate significantly less often than unrelated males and gynes." However, *P. versicolor* males did not avoid relatives or discriminate between female castes in laboratory mate choice assays (de Souza et al. 2017a; de Souza et al. 2021). Similarly, there is evidence for (Shellman-Reeve & Gamboa 1985) and against (Ryan et al. 1984) the claim that *P. fuscatus* males can recognize male nestmates. In *Polistes*, nestmate recognition phenotypes are complex cuticular hydrocarbon profiles (Gamboa et al. 1986; Espelie et al. 1994; Dani et al. 2001). Future work should investigate how odorant receptors may mediate nestmate recognition and other forms of olfactory recognition in *P. fuscatus*.

Acknowledgments

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.], and National Institutes of Health [grant number DP2-GM128202 to M.J.S.].

REFERENCES

Alexa A, Rahnenfuhrer J. 2021. topGO: enrichment analysis for gene ontology. R package version 2.46.0.

Beani L, Turillazzi S. 1988. Alternative mating tactics in males of *Polistes dominulus* (Hymenoptera: Vespidae). Behavioral Ecology and Sociobiology. 22:257-264.

Beani L, Dessì-Fulgheri F, Cappa F, Toth A. 2014. The trap of sex in social insects: from the female to the male perspective. Neuroscience and Biobehavioral Reviews. 46(4):519-533.

Beani L, Bagnères A-G, Elia M, Petrocelli I, Cappa F, Lorenzi MC. 2019. Cuticular hydrocabons as cues of sex and health condition in *Polistes dominula* wasps. Insectes Sociaux.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B. 57:289-300.

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

Berson JD, Zuk M, Simmons LW. 2019. Natural and sexual selection on cuticular hydrocarbons: a quantitative genetic analysis. Proceedings of the Royal Society B. 286:20190677.

Bluher SE, Miller SE, Sheehan MJ. 2020. Fine-scale population structure but limited genetic differentiation in a cooperatively breeding paper wasp. Genome Biol Evol. 12:701-714.

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

Boomsma JJ, Kronauer DJC, Pedersen JS. 2009. The Evolution of Social Insect Mating Systems. In Organization of Insect Societies. Eds Gadau J, Fewell J. Harvard University Press. Pages 3-25.

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 Ecology and Evolution. 15:176.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 25(15):1972-1973.

Cappa F, Bruschini C, Cervo R, Turillazzi S, Beani L. 2013. Males do not like the working class: male sexual preference and recognition of functional castes in a primitively eusocial wasp. Animal Behaviour. 86:801-810.

Chen Y, McCarthy D, Ritchie M, Robinson M, Smyth G. 2008. edgeR: differential expression analysis of digital gene expression data User's Guide.

da Silva RC, Meerbeeck LV, do Nascimento FS, Wenseleers T, Oi CA. 2021. Close-range cues used by males of *Polistes dominula* in sex discrimination. The Science of Nature. 108:15

Dani FR, Jones GR, Destri S, Spencer SH, Turillazzi S. 2000. Deciphering the recognition signature within the cuticular chemical profile of paper wasps. Animal Behaviour. 62(1):165-171.

de Souza AR, Barbosa BC, da Silva RC, Prezoto F, Lino-Neto J, do Nascimento FS. 2017a. No evidence of intersexual kin recognition by males of the neotropical paper wasp *Polistes versicolor*. Journal of Insect Behavior. 30:180-187.

de Souza AR, Lino-Neto J, Tibbetts EA, Turillazzi S, Beani L. 2017b. The leks of *Polistes dominula* paper wasps: tiny abdominal spots play a critical role in male attacks toward potential rivals. Ethology Ecology and Evolution. 29(4):410-419.

de Souza AR, da Silva RC, Batista NR, Júnior WFA, do Nascimento FS. 2021. Lack of caste discrimination by males during sexual context in a neotropical paper wasp. Ethology. 127(7):613-619.

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.

Durand N, Carot-Sans G, Bozzolan F, Rosell G, Siaussat D, Debernard S, Chertemps T, Maïbèche-Coisne M. 2011. Degradation of pheromone and plant volatile components by a same odorant-degrading enzyme in the cotton leafworm, *Spodoptera littoralis*. PLOS ONE. 6(12):e29147.

Elmquist DC, Landolt PJ, Ream LM, Cha DH. 2018. Laboratory demonstrations of pheromone-mediated scent-marking, orientation, and mounting behavior in *Polistes exclamans* (Hymenoptera: Vespidae). Annals of the Entomological Society of America. 111:21-30.

Espelie KE, Hermann HR. 1990. Surface lipids of the social wasp *Polistes annularis* (L.) and its nest and nest pedicel. Journal of Chemical Ecolgy. 16:1841-1852.

Espelie KE, Gamboa GJ, Grudzien TA, and Bura EA. 1994. Cuticular hydrocarbons of the paper wasp *Polistes fuscatus*: a search for recognition pheromones. Journal of Chemical Ecology 20(7):1677-1687.

Feyereisen R. 2006. Evolution of insect P450. Biochemical Society Transactions. 34(6):1252-1255.

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

Hunt JH, Noonan KC. 1979. Larval feeding by male *Polistes fuscatus* and *Polistes metricus* (Hymenoptera: Vespidae). Insectes Sociaux. 26(3):247-251.

Ishida Y, Leal WS. 2005. Rapid inactivation of a moth pheromone. Proceedings of the National Academy of Sciences USA. 102(39):14075-14079.

Ishida Y, Leal WS. 2008. Chiral discrimination of the Japanese beetle sex pheromone and a behavioral antagonist by a pheromone-degrading enzyme. Proceedings of the National Academy of Sciences USA. 105(26):9076-9080.

Izzo AS, Tibbetts EA. 2012. Spotting the top male: sexually selected signals in male *Polistes dominulus* wasps. Animal Behaviour. 83:839-845.

Jernigan CM, Zaba NC, Sheehan MJ. 2021. Age and social experience induced plasticity across brain regions of the paper wasp *Polistes fuscatus*. Biology Letters. 17(4):20210073.

Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Bioinformatics. 8(3):275-282.

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 *Apis florea* odorant receptors reveals complex orthologous relationships with *Apis mellifera*. Genome Biol Evol. 8(9):2879–2895.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution. 30:772-780.

Koch SI, Groh K, Vogel H, Hansson BS, Kleineidam CJ, Grosse-Wilde E. 2013. Caste-specific expression patterns of immune response and chemosensory related genes in the leaf-cutting ant, *Atta vollenweideri*. PLoS ONE 8:e81518.

Koeniger N, Koeniger G, Gries M, Tingek S. 2005. Drone competition at drone congregation areas in four *Apis* species. Apidologie. 36(2):211-221.

Lande R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. Evolution 34:292-305.

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 9:559.

Larch CM, Gamboa GJ. 1981. Investigation of mating preference for nestmates in the paper wasp Polistes fuscatus (Hymenoptera: Vespidae). J. Kans. Entomol. Soc. 54:811-814.

Layton JM, Camann MA, Espelie KE. 1994. Cuticular lipid profiles of queens, workers, and males of social wasp *Polistes metricus* say are colony-specific. Journal of Chemical Ecology. 20:2307-2321.

Leal WS. 2013. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. Annual Review of Entomology. 58:373-391.

Legan AW, Jernigan CM, Miller SE, Fuchs MF, Sheehan MJ. 2021. Expansion and accelerated evolution of 9-exon odorant receptors in *Polistes* paper wasps. Molecular Biology and Evolution 38(9):3832-3846.

McKenzie SK, Oxley PR, Kronauer DJC. 2014. Comparative genomics and transcriptomics in ants provide new insights into the evolution and function of odorant binding and chemosensory proteins. BMC Genomics. 15:718.

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. Proceedings of the National Academy of Sciences 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 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 MT, Ostevik KL, Samuk K, Uy FMK, Sheehan MJ. 2020. Evolutionary dynamics of recent selection on cognitive abilities. Proceedings of the National Academy of Sciences USA. 117(6):3045-3052.

Miller SE, Sheehan MJ. 2021. Ecogeographical patterns of body size differ among North American paper wasp species. Insectes Sociaux. 68:109-122.

Nakagawa T, Sakurai T, Nishioka T, Touhara K. 2005. Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. Science 307(5715):1638-1642.

Nakanishi A, Nishino H, Watanabe H, Yokohari F, Nishikawa M. 2009. Sex-specific antennal sensory system in the ant *Camponotus japonicus*: structure and distribution of sensilla on the flagellum. Cell and Tissue Research. 338:79-97.

O'Donnell S, Clifford MR, Bulova SJ, DeLeon S, Papa C, Zahedi N. 2014. A test of neuroecological predictions using paperwasp caste differences in brain structure (Hymenoptera: Vespidae). Behavioral Ecology and Sociobiology. 68:529-536.

Parsch J, Ellegren H. 2013. The evolutionary causes and consequences of sex-biased gene expression. Nature Reviews Genetics 14:83-87.

Pekkarinen A, Gustafsson B. 1999. The *Polistes* species in northern Europe (Hymenoptera: Vespidae). Entomologica Fennica. 10:191-197.

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

Post DC, Jeanne RL. 1983b. Relatedness and mate selection in Polistes fuscatus (Hymenoptera: Vespidae). Animal Behaviour 31:1260-1261.

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.

Qiu Y, Tittiger C Wicker-Thomas C, Le Goff G, Young S, Wajnberg E, Fricaux T, Taquet N, Blomquist GJ, Feyereisen R. 2012. An insect-specific P450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis. Proceedings of the National Academy of Sciences USA. 109(37):14858-14863.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26:841-842.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/.

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.

Rihani K, Sachse S. 2022. Shedding light on inter-individual variability of olfactory circuits in *Drosophila*. Frontiers in Behavioral Neuroscience. 16:835680:1-11.

Robertson HM, Warr CG, Carlson JR. 2003. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA. 100(suppl_2):14537-14542.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 26:139-140.

Romani R, Isidoro N, Riolo P, Bin F, Fortunato A, Turillazzi S, Beani L. 2005. A new role for antennation in paper wasps (Hymenoptera, Vespidae): antennal courtship and sex dimorphic glands in antennomeres. Insectes Sociaux. 52:96-102.

Rospars JP, Chambille I. 1989. Identified glomeruli in the antennal lobes of insects: in variance, sexual variation and postembryonic development. In: Singh RN, Strausfeld NJ (eds) Neurobiology of Sensory Systems. Springer, Boston, MA.

Ryan RE, Forbes GC, Gamboa GJ. 1984. Male social wasps fail to recognize their brother (*Polistes fuscatus*, Hymenoptera: Vespidae). Journal of the Kansas Entomological Society. 57(1):105-110.

Ryan RE, Gamboa GJ. 1986. Nestmate recognition between males and gynes of the social wasp Polistes fuscatus (Hymenoptera: Vespidae). Annals of the Entomological Society of America 79:572-575.

Scherberich J, Hummel J, Schöneich S, Nowotny M. 2017. Functional basis of the sexual dimorphism in the auditory fovea of the duetting bushcricket *Ancylecha fenestrata*. Proceedings of the Royal Society B. 284:20171426.

Shellman-Reeve J, Gamboa GJ. 1985. Male social wasps (Polistes fuscatus, hymenoptera: Vespidae) recognize their male nestmates. Animal Behaviour 33:331-333.

Shigenobu S, Hayashi Y, Watanabe D, Tokuda G, Hojo MY, Toga K, Saiki R, Yaguchi H, Masuoka Y, Suzuki R, et al. 2022. Genomic and transcriptomic analyses of the subterranean termite Reticulitermes speratus: gene duplication facilitates social evolution. Proceedings of the National Academy of Sciences USA. 119(3):e2110361119.

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. Proceedings of the National Academy of Sciences USA. 114(32):8586-8591.

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*). Proceedings of the National Academy of Sciences USA. 108(14):5673-5678.

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*. Proceedings of the National Academy of Sciences USA. 108(14):5667-5672.

Somanathan H, Borges RM, Warrant EJ, Kelber A. 2017. Visual adaptations for mate detection in the male carpenter bee *Xylocopa tenuiscapa*. PLOS ONE 12(1):e0168452.

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

Suh E, Bohbot J, Zwiebel LJ. 2014. Peripheral olfactory signaling in insects. Current Opinion in Insect Science. 6:86-92.

Sutherland TD, Unnithan GC, Andersen JF, Evans PH, Murataliev MB, Szabo LZ, Mash EA, Bowers WS, Feyereisen R. 1998. A cytochrome P450 terpenoid hydroxylase linked to the suppression of insect juvenile hormone synthesis. Proceedings of the National Academy of Sciences USA. 95:12884-12889.

Uy FMK, Jernigan CM, Zaba NC, Mehrotra E, Miller SE, Sheehan MJ. 2021. Dynamic neurogenomic responses to social interactions and dominance outcomes in female paper wasps. PLOS Genetics. 17(9):e1009474.

Wanner KW, Anderson AR, Trowell SC, Theilmann DA, Robertson HM, Newcomb RD. 2007. Female-biased expression of odourant receptor genes in the adult antennae of the silkworm, *Bombyx mori*. Insect Molecular Biology. 16(1):107-119.

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. Proceedings of the National Academy of Sciences USA 104(36):14383-14388.

West-Eberhard MJ. 1969. The social biology of Polistine wasps. Miscellaneous publications of the museum of zoology, University of Michigan. 140:1-110.

Yu G, Smith D, Zhu H, Guan Y, Lam TT-Y. 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution. 8:28-36.

Zhao X-C, Ma B-W, Berg BG, Xie G-Y, Tang Q-B, Guo X-R. 2016. A global-wide search for sexual dimorphism of glomeruli in the antennal lobe of female and male *Helicoverpa armigera*. Scientific Reports. 6:35204.

Zhao Z, McBride CS. 2020. Evolution of olfactory circuits in insects. Journal of Comparative Physiology A. 206:353-367.

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.

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 Biology and Evolution. 7(8):2407-2416.

CHAPTER 4

POSTURAL ANALYSIS REVEALS PERSISTENT VIGILANCE IN PAPER WASPS AFTER CONSPECIFIC CHALLENGE

Andrew W. Legan, Caleb C. Vogt, Michael J. Sheehan

Preprint posted on bioRxiv May 29, 2022, DOI: 10.1101/2022.05.25.493496

Abstract

Vigilant animals detect and respond to threats in the environment, often changing posture and movement patterns. In social animals vigilance is modulated not only by predators but also by threatening conspecifics. Precisely how social interactions alter vigilance behavior over time is not well understood. We report persistent effects of a simulated social challenge on the vigilance behavior in wild northern paper wasp foundresses, *Polistes fuscatus*. During the founding phase of the colony cycle conspecific wasps may usurp nests from the resident foundress, representing a severe threat. Using postural tracking, we found that after simulated intrusions wasps displayed increased vigilance during the minutes after the threat was removed. Sustained vigilance elicited after social threat manifested as increased movement, greater bilateral wing extension, and reduced antennal separation. However, no postural changes were observed after a control stimulus presentation. By rapidly adjusting individual vigilance behavior after fending off a conspecific intruder, paper wasp foundresses invest in surveillance of potential social threats, even when such threats are no longer immediately present. The prolonged state of vigilance observed here is relevant to plasticity of recognition processes as a result of conspecific threats.

Introduction

Vigilance behavior in animals is demonstrated by changes in movement and body posture, famously in the still, bipedal stance of meerkat sentinels (Santema and Clutton-Brock 2013). Movement and posture of specific body parts, especially the head and sensory organs, are primarily responsible for vigilance quality because they directly influence perception. For example, chaffinches turn their heads more after seeing a cat (Jones et al. 2007) and vigilant baboons blink less (Matsumoto-Oda et al. 2018). Animals must sometimes sacrifice vigilance quality in favor of other important activities, for example in the case of feeding juncos forfeiting some vigilance quality to lower their heads and eat (Lima and Bednekoff 1999).

Social animals, though characterized by their cooperative associations, face threats posed by conspecifics (Abbot 2022). Recognition is an important mechanism mediating intraspecific aggression because encounters with different individuals and classes of individuals may impact fitness in distinct ways (Mateo 2004; Bourke 2011; Gherardi et al. 2012; Leonhardt et al. 2016; Sheehan and Bergman 2016). Social insects exhibit plasticity in nest guarding behavior in response to the frequency and valence of interactions with different classes of individuals (e.g., nestmates and non-nestmates) (Starks et al. 1998; Liebert and Starks 2004; Fürst et al. 2011; Mora-Kepfer 2014). In response to encounters with non-nestmates, honeybees restrict admittance to the colony, sometimes rejecting their own nestmates (Downs and Ratnieks 2000; Couvillon et al. 2008). These rejection errors are consistent with the signal detection theory concept of a shifting acceptance threshold (Reeve 1989). With more frequent intruder encounters, the cost of accidentally accepting intruders increases, and the acceptance

threshold is reduced to minimize acceptance errors. An alternative view considers variation in recognition behavior in terms of investment in recognition accuracy (Sheehan and Reeve 2020). Recognition accuracy may be improved by persistent vigilant behavior of guards. Shifts in vigilance at the group level have been documented in honey bees, which allocate more guards at the colony entrance in response to threats (Downs and Ratnieks 2000; Breed et al. 1992). How persistent vigilance is manifested in individual movement and posture has not been examined in a social insect.

To approach this question, we studied the northern paper wasp *Polistes fuscatus*. Paper wasps are ideal for field-based digital tracking because their unenveloped nest represents a fixed arena easily recordable by video. Automated tracking of wild foundress behavior is an as-yet unapplied tool for understanding the effects of intruder encounters on vigilance. During nest founding in the spring, *Polistes* foundresses guard the nest from conspecific wasps which may rob their brood or usurp their nests (Reeve 1991; Sheehan et al. 2015; Miller et al. 2018; Sakagami and Fukushimna 1957; Kasuya et al. 1980; Gamboa et al. 1992). We simulated a guard context during the founding phase of single foundress *P. fuscatus* nests and leveraged digital tracking software to analyze wasp movement and posture.

Methods

We studied solitary *P. fuscatus* foundresses on their nests at the Liddell Field Station in Ithaca, NY (42°27’36.7” N, 76°26’39.2” W). In the spring of 2020, wild wasps initiated nests in modified wooden bird boxes (11.5 cm x 12.5 cm x 13.5 cm). All experiments were carried out from July 4th to July 9th, 2020, before workers emerged. Experiments were carried out between 2PM and 8PM EST, during the active

phase of wasps during peak summer. The mean nest size was 33 ± 8 (SD) cells. The experimental apparatus consisted of a 162.5 cm wooden dowel (7 mm diameter) guided through a 122 cm metal cylinder (1 cm diameter), taped to a step ladder (fig. 4.1). The assays were video-recorded from below using a tripod-mounted Nikon D7200 camera with a Sigma Macro HSM lens with an optical stabilizer (focal length: 105 mm; aperture: f/2.8).

Figure 4.1. Experimental apparatus.

Intruder wasps were collected from nests at a site (42°24'57.6" N, 76°31'22.6" W) 8.15 km southwest of the Liddell Station to ensure that foundresses had not encountered intruders before the experiment and were unlikely to be closely related (Bluher et al. 2020). Wasps were size matched to lures within 0.028 ± 0.013 grams (SD). Immediately before each simulated intruder trial, the intruder wasp was freeze-killed and fixed to a wooden dowel using an insect pin. Unique intruders were presented as the stimulus in each simulated intruder trial. On a different day, each guard wasp was presented with the wooden dowel alone. All assays consisted of three 320 second intervals: pre-stimulus, stimulus, and post-stimulus. All nests were undisturbed, with experimental apparatus in place, for at least 5 minutes before beginning the pre-

stimulus interval. During the stimulus presentation in both simulated intruder and wooden dowel trials, the stimulus was moved slightly by the experimenter at one-minute intervals to animate the stimulus. Three foundresses were excluded from analysis because a live intruder visited the nest during the experiment and one foundress was excluded from analysis because the foundress was accidentally flushed from the nest while setting up the experimental apparatus. Ultimately six foundresses were assayed.

We used computer vision software SLEAP (Pereira et al. 2022) to track seven points on the wasps: antennae tips, head, thorax-abdomen bridge (propodeum), abdomen tip, and wing tips (fig. 4.2A). SLEAP was installed on a PC equipped with a GeForce RTX 2080i graphics card. Videos were converted to gray scale and a subset of 20 frames per interval were manually labeled. Output files and tracked video files are available online in the supplement and on Zenodo (https://zenodo.org/record/6582229#.Yo7Of). We compared wing and antennae separation angle before and after stimulus presentations using paired t-tests. Statistical analysis was done using R version 3.6.1 (R Core Team 2019).

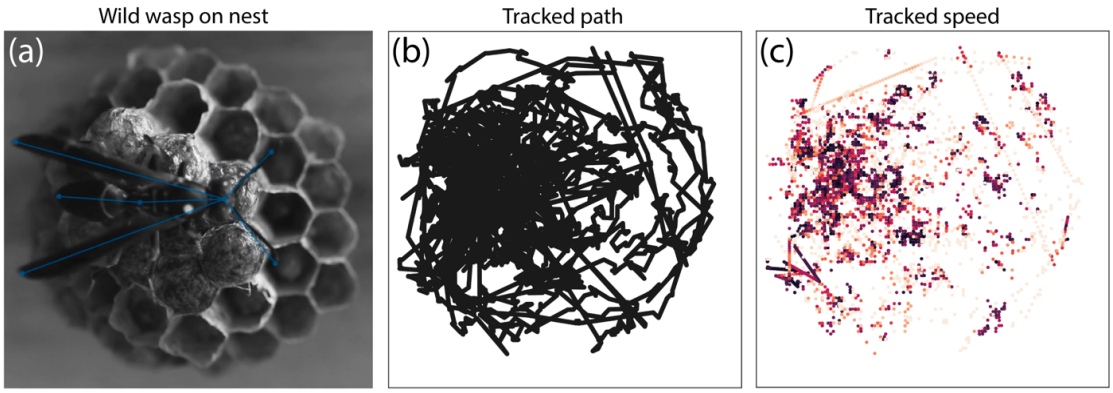

Figure 4.2. (a) A lone *Polistes fuscatus* foundress on the nest after a simulated intruder encounter. (b) Tracks of the position of the thorax of the wasp over a 320 second

interval after simulated intrusion. (c) Points designate the position of the thorax and are color-coded by speed, with lighter color representing faster movement.

Results

When the lure was presented during simulated intruder trials, wasps responded by antennating the lure, then responded aggressively by biting, mounting, and stinging the lure. These are all stereotyped wasp aggressive behaviors (West-Eberhard 1969; Lorenzi et al. 1997; Tumulty et al. 2021). In control stimulus trials, wasps investigated the dowel, including antennation and occasional mounting, but did not escalate aggression. SLEAP successfully tracked body parts in 84 ± 21% (SD) of frames across body parts before and after stimulus presentation (table S4.1).

Simulated intruder encounters caused persistent changes in posture while control experiments did not. Exposure to the simulated intruder caused an increase in the total distance traveled after the intruder was removed (head: $t = -2.5682$, $df = 5$, P-value = 0.05014; thorax: $t = -2.9614$, $df = 5$, P-value = 0.03147; fig. 4.3A). Dowel presentations did not lead to sustained increases in movement afterward (head: $t = 0.27264$, $df = 5$, P-value = 0.796; thorax: $t = 0.037249$, $df = 5$, P-value = 0.9717). Wing posture was affected by the simulated intruder. The mean wing extension angle after intruder encounter was significantly greater than before ($t = -6.1917$, $df = 5$, P-value = 0.001603; fig. 4.3B). No significant change in wing extension angle was observed after the wooden dowel presentation ($t = -1.2836$, $df = 5$, P-value = 0.2556). There was a significant decrease in the mean antennal separation angle after intruder encounter ($t = 4.1753$, $df = 5$, P-value = 0.008695; fig. 4.3C). No significant change in mean antennal separation angle was observed after the wooden dowel presentation ($t = 0.40974$, $df = 5$, P-value = 0.699).

The rapid and varied movement of wasps during simulated intruder presentations, and the presence of a second, pinned wasp, precluded successful digital tracking. However, digital tracking during the dowel presentations was feasible. During the dowel presentation, wasps did not move more than they did before the presentation, based on the total distance traveled by the thorax ($t = -1.2475$, $df = 5$, P-value = 0.2675; fig. 4.4A). There was a significant increase in wing extension angle during the dowel presentation compared to before ($t = -2.8063$, $df = 5$, P-value = 0.03771; fig. 4.4B). This increase in wing extension did not persist after the dowel was removed.

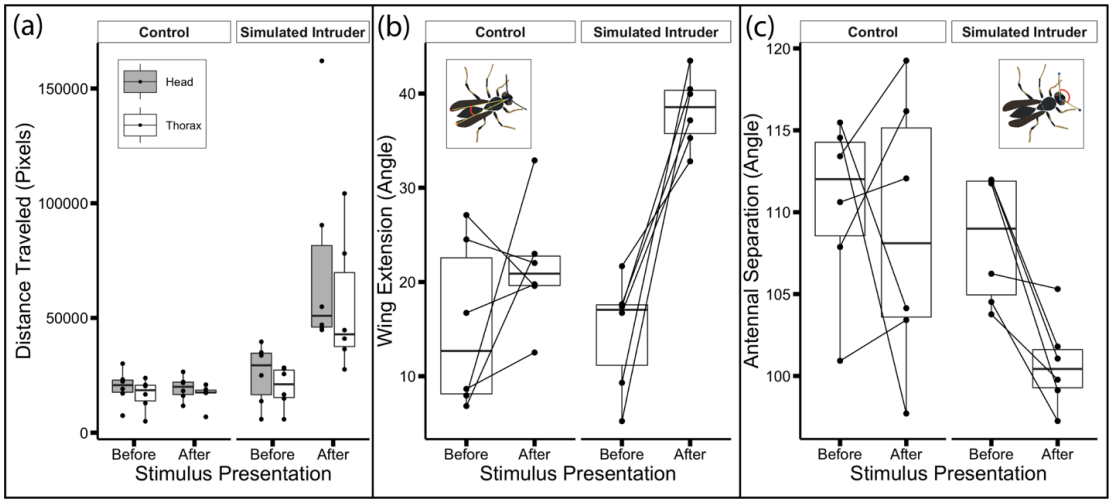

Figure 4.3. Box and whisker plots display comparisons of measures of movement and posture across trials. (a) Total distance traveled by head (gray) and thorax. (b) Wing extension angle. (c) Antennal separation angle.

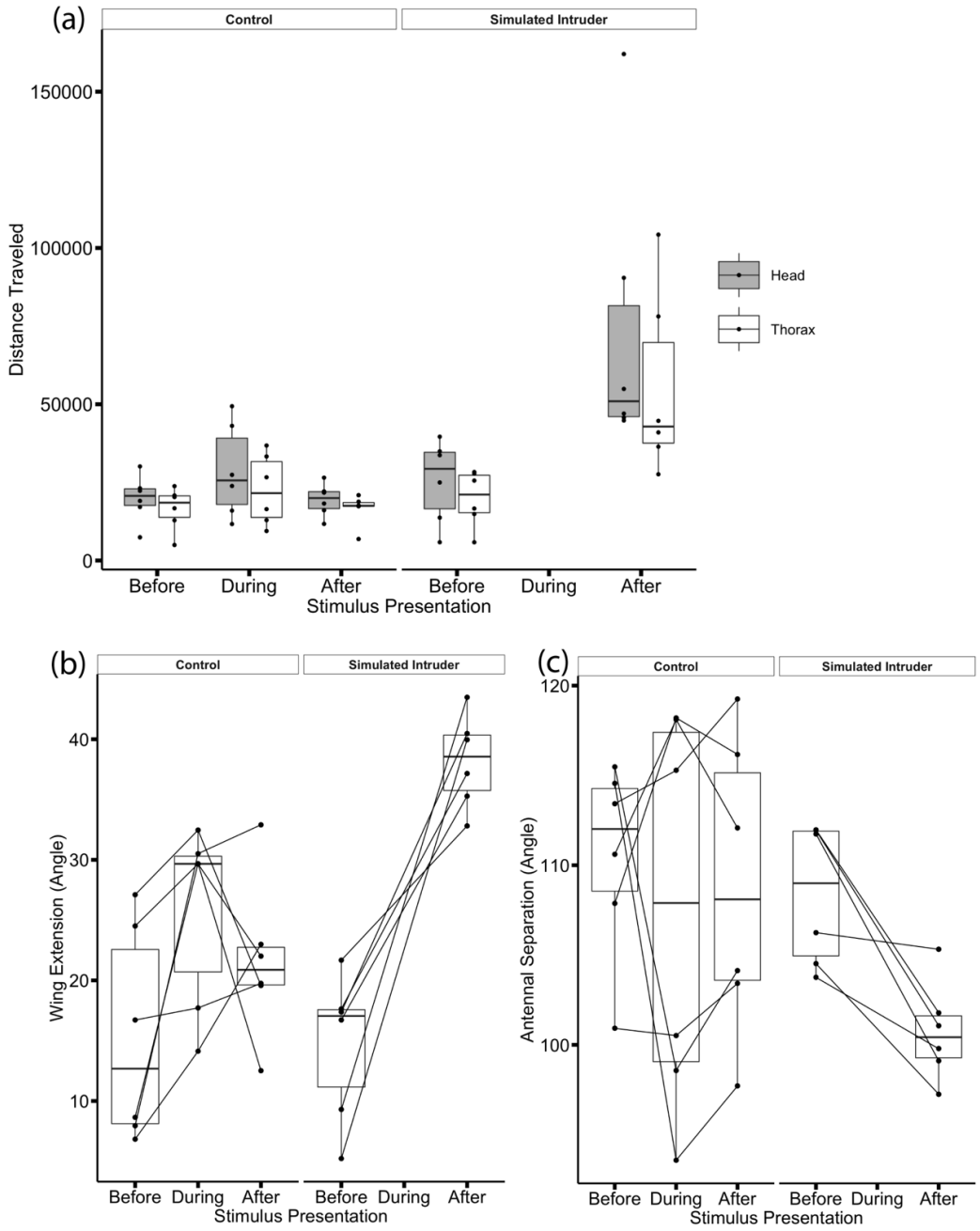

Figure 4.4. Box and whisker plots display comparisons of measures of movement and posture across trials. Same data as fig. 4.3, except interval 2 of control trials has been tracked. (a) Total distance traveled by head (gray) and thorax in each interval. (b) Wing extension measured as an angle in degrees. (c) Antennal separation measured as an angle in degrees. In each plot, the horizontal line marks the median of the data points, the box brackets the middle fifty percent of data points, and the whiskers extend to the smallest and largest values at most 1.5 * the interquartile range.

Discussion

Social challenge presented by simulated conspecific intruders elicited sustained vigilance in *P. fuscatus* spring foundresses. We leveraged computer vision to analyze wasp body posture in the field and found that sustained vigilance manifested in changes to wasp movement and posture.

Methods in automated tracking of behavior have been applied by scientists studying neurobiological mechanisms of animal movement and pose, collective behavior, and social interactions (Dell et al. 2014; Crall et al. 2018; Mathis and Mathis 2020; Wang et al. 2022). These studies are often carried out in controlled environments, which is feasible when the behavior of interest is robust to laboratory conditions. For example, digital tracking has been used to characterize the foraging behavior of hawkmoths *Manduca sexta* (Dahake et al. 2018; Deora et al. 2021), and to characterize the wing kinematics of flies and honey bees as well as honey bee wing fanning behavior (Altshuler et al. 2005; Muijres et al. 2014; Peters et al. 2017). Complex social behaviors are less robust to experimental laboratory conditions, requiring field observations to draw reliable conclusions. But few studies have applied digital tracking of individual social animal posture in the wild (but see Peters et al. 2017).

Polistes paper wasps are ideal for computer vision-assisted digital tracking and pose estimation in the field. Compared to eusocial ants, bees, and hornets, *Polistes* societies remain relatively small, peaking at ~135 cells (Reeve 1991). *Polistes* colonies are generally single-layer nests. In terms of video recording, a drawback to this architecture is that there is usually space between the nest and the substrate to which it is fixed, so wasps can crawl out of view of the camera behind the nest. While the nest

can be treated as two-dimensional for the purpose of digital tracking, the wasp's body is not always parallel to this plane, leading to difficulties in tracking a wasp perched on the side of the nest. These challenges could be solved in principle with multiple cameras recording the nest from different angles, as has been done recently for 3-dimensional tracking in laboratory rodents (Marshall et al. 2021; Ebbesen and Froemke 2022). Another challenge for digital tracking is the rapid movement of wasps during the simulated intrusions, but cameras with faster frame rates could overcome this issue.

Natural threats that would induce nest-guarding behavior in solitary foundresses include intraspecific brood-robbing and nest usurpation (Sakagami and Fukushima 1957, Kasuya et al. 1980; Gamboa et al. 1992). In a study of multiple foundress *P. fuscatus* nests during the founding phase, natural encounters with intruders occurred about once per day, with intruders evicted within 40 seconds (Gamboa et al. 1992). Three trials in our study were interrupted by natural intruders, highlighting the pervasive nature of conspecific threats for *P. fuscatus* foundresses. The 320 second lure presentation in our assays likely simulated a worst-case scenario for foundresses, akin to a nest usurpation attempt.

Postural changes displayed by vigilant wasps included wing extension and reduced antennal separation. Upon presentation with the simulated intruder, wasps approached the lure with outstretched antennae before reacting aggressively. In general, social insects utilize chemical cues to discriminate nestmates and non-nestmates (Nunes et al. 2008; Van Zweden and d'Ettorre 2010). While *P. fuscatus* wasps rely on vision to recognize individuals, nestmate recognition is mediated by olfaction (Tibbetts 2002; Ortiz and Tibbetts 2020). Reduced antennal separation may indicate that wasps are

orienting their antennae to detect chemical cues, such as the cuticular hydrocarbon signatures used by social insects to discriminate nestmates and non-nestmates (Gamboa et al. 1986; Dani et al. 2001; Bruschini et al. 2011; Nascimento and Nascimento 2012). Visual cues may also be important in discriminating nestmates and non-nestmates at the early phases of the colony cycle, and the absence of nestmates may favor universal rejection (Reeve 1989; Cini et al. 2019).

Vigilant wasps moved more after fighting a simulated intruder, as measured by total distance traveled. This increased movement was observed throughout the 320 s interval after the simulated intruder was removed (fig. 4.5). By moving throughout the nest surface, vigilant guard wasps may be better prepared to defend against an intruder approaching from any direction.

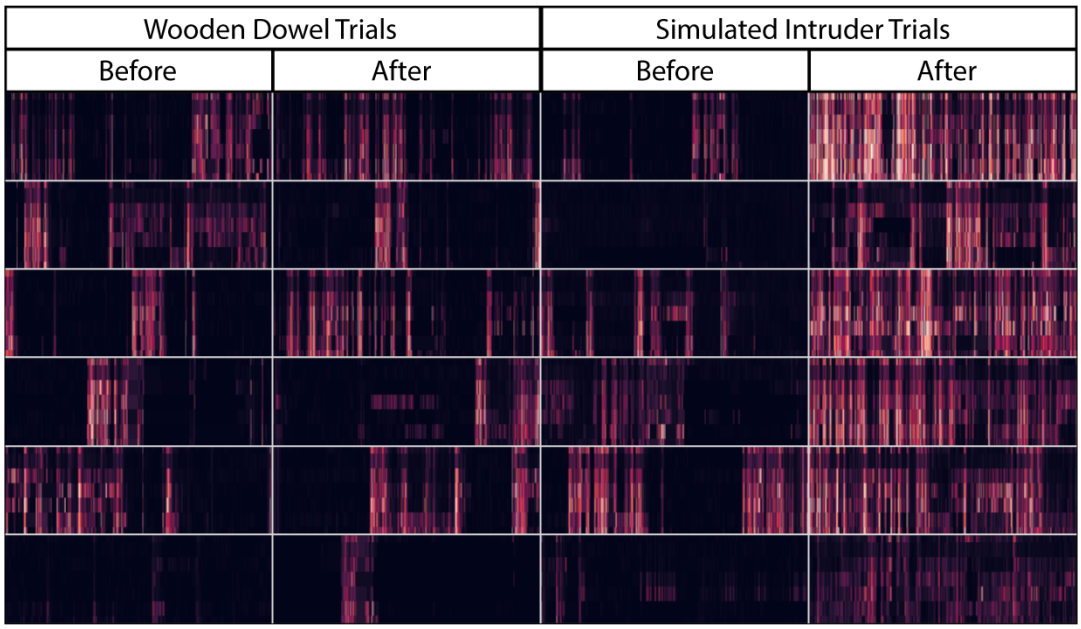

Figure 4.5. For six wasps assayed in two experimental intervals of twelve trials, the speeds of seven tracked body parts over time are shown in a heatmap. Six rows of heatmaps correspond to six individual wasps presented with a wooden dowel and an intruder wasp on separate days. Rows within each heatmap correspond, from top to bottom, to seven tracked body parts: head, thorax-abdomen bridge (propodeum), abdomen tip, left wing tip, right wing tip, left antenna tip, and right antenna tip.

In *Polistes*, wing extension and antennal separation may be useful measures for studying how the social environment influences internal state. The reliable associations between unilateral wing-extension and courtship, and between bilateral wing extension and aggression, have been useful measures for studying the neural basis of aggression and courtship in the fly *Drosophila melanogaster*, especially the roles of P1 neurons in orchestrating persistent internal states causing aggression and courtship (Zhou et al. 2008; Hoopfer et al. 2015).

The internal state associated with vigilance in *P. fuscatus* may represent an emotional primitive, as defined by Anderson and Adolphs (Anderson and Adolphs 2014) as an internal state exhibiting scalability, valence, persistence, and generalization. Regarding scalability, we found preliminary evidence that wing extension can be ordered along a gradient corresponding to low vigilance (before stimulus), medium vigilance (during dowel presentation), and high vigilance (after simulated intruder presentation, demonstrating behavioral persistence) (fig. 4.4). *P. fuscatus* vigilance behavior is associated with aggression towards conspecific intruders, suggesting negative valence. After social challenge, vigilance is persistent. More work needs to be done to assess the generalization of *P. fuscatus* vigilance behavior, for example by presenting wasps with neutral stimuli after social challenge.

Increased encounters with non-nestmate intruders can shift social insect recognition processes to become more exclusive, resulting in recognition error in the form of increased aggression towards nestmates (Starks et al. 1998; Mora-Kepfer 2014; Downs and Ratnieks 2000; Couvillon et al. 2008; Scharf et al. 2020). From the perspective of signal detection theory, individual vigilance behavior could be

mechanistically related to acceptance threshold shifts. If persistent vigilance and acceptance threshold shift are coupled, then there will be more aggression towards nestmates following intruder encounters. Alternatively, persistent vigilance may have effects on recognition independent of acceptance threshold shifts. For example, persistent vigilance may accompany increased investment in accurate recognition (Sheehan and Reeve 2020). Evidence supporting this alternative may be found in the carpenter ant, where exposure to alarm pheromone increased accuracy of both nestmate acceptance and non-nestmate rejection (Rossi et al. 2019). Persistent vigilance may therefore increase recognition accuracy, while the acceptance threshold is shifted depending on non-nestmate encounter rates (Reeve 1989). Future work should explore how individual wasp vigilance behavior relates to shifts in nestmate recognition processes.

Acknowledgments

This work was supported by National Science Foundation [Graduate Research Fellowship Program DGE-1650441 to AWL, CAREER grant DEB-1750394 to MJS], Cornell University [Neurobiology and Behavior Departmental Grant to CCV], and National Institutes of Health [DP2-GM128202 to MJS], and by the North American Section of the International Union for the Study of Social Insects (IUSSI) Robert L. and Louise B. Jeanne Social Wasp Research Grant [to AWL].

REFERENCES

Santema P, Clutton-Brock T. (2013) Meerkat helpers increase sentinel behaviour and bipedal vigilance in the presence of pups. Animal Behaviour 85(3):655-661.

Jones KA, Krebs JR, Whittingham MJ. (2007) Vigilance in the third dimension: head movement not scan duration varies in response to different predator models. Animal Behaviour 74(5):1181-1187.

Matsumoto-Oda A, Okamoto K, Takahashi K, Ohira H. (2018) Group size effects on inter-blink interval as an indicator of antipredator vigilance in wild baboons. Scientific Reports 8:10062.

Lima SL, Bednekoff PA. 1999. Back to the basics of antipredatory vigilance: can nonvigilant animals detect attack? Animal Behaviour 58:537-543.

Abbot P. (2022) Defense in social insects: diversity, division of labor, and evolution. Annual Review of Entomology 67:407-436.

Mateo JM. (2004) Recognition systems and biological organization: the perception component of social recognition. Ann. Zool. Fennici 41:729-745.

Bourke AFG. (2011) Social group maintenance. In Principles of Social Evolution. Oxford University Press Inc., New York.

Gherardi F, Aquiloni L, Tricarico E. (2012) Revisiting social recognition systems in invertebrates. Animal Cognition 15(5):745-762.

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

Sheehan MJ, Bergman TJ. 2016. Is there an evolutionary trade-off between quality signaling and social recognition? Behavioral Ecology 27(1):2-13.

Starks PT, Fischer DJ, Watson RE, Melikian GL, Nath SD. (1998) Context-dependent nestmate discrimination in the paper wasp, *Polistes dominulus*: a critical test of the optimal acceptance threshold model. Animal Behaviour 56:449-458.

Liebert AE, Starks PT. (2004) The action component of recognition systems: a focus on the response. Annales Zoologici Fennici 41(6):747-764.

Fürst MA, Durey M, Nash DR. (2011) Testing the adjustable threshold model for intruder recognition on *Myrmica* ants in the context of a social parasite. Proceedings of the Royal Society B 279:516-522.

Mora-Kepfer F. (2014) Context-dependent acceptance of non-nestmates in a primitively eusocial insect. Behavioral Ecology and Sociobiology 68:363-371.

Downs SG, Ratnieks FLW. (2000) Adaptive shifts in honey bee (*Apis mellifera* L.) guarding behavior support predictions of the acceptance threshold model. Behavioral Ecology 11(3):326-333.

Couvillon MJ, Robinson EJH, Atkinson B, Child L, Dent KR, Ratnieks FLW. (2008) En garde: rapid shifts in honeybee, *Apis mellifera*, guarding behaviour are triggered by onslaught of conspecific intruders. Animal Behaviour 76(5):1653-1658.

Reeve HK. (1989) The evolution of conspecific acceptance thresholds. The American Naturalist 133:407-435.

Sheehan MJ, Reeve HK. (2020) Evolutionarily stable investments in recognition systems explain patterns of discrimination failure and success. Philosophical Transactions of the Royal Society B. 375(1802): 20190465.

Breed MD, Smith TA, Torres A. (1992) Role of guard honey bees (Hymenoptera: Apidae) in nestmate discrimination and replacement of removed guards. Annals of the Entomological Society of America 85(5):633-637.

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

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. Ecology Letters 18:1057-1067.

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.

Sakagami SF, Fukushima K. (1957) Reciprocal thieving found in *Polistes fadwigae* Dalla Torre. J. Kans. Entomol. Soc. 30:140.

Kasuya E, Hibino Y, Itô Y. (1980) On "intercolonial" cannibalism in Japanese paper wasps, *Polistes chinensis antennalis* Pérez and *P. jadwigae* Dalla Torre (Hymenoptera: Vespidae). Res. Popul. Ecol. 22:255-262.

Gamboa GJ, Wacker TL, Duffy KG, Dobson SW, Fishwild TG. (1992) Defence against intraspecific usurpation by paper wasp cofoundresses (*Polistes fuscatus*, Hymenoptera: Vespidae). Can. J. Zool. 70:2369-2372.

Bluher SE, Miller SE, Sheehan MJ. (2020) Fine-scale population structure but limited genetic differentiation in a cooperatively breeding paper wasp. Genome Biology and Evolution 12(5):701-714.

Pereira TD, Tabris N, Matsliah A, Turner DM, Li J, Ravindranath S, Papadoyannis ES, Normand E, Deutsch DS, Wang ZY, McKenzie-Smith GC, Mitelut CC, Castro MD, D'Uva J, Kislin M, Sanes DH, Kocher SD, Wang SSH, Falkner AL, Shaevitz JW, and Murthy M. (2022) Sleap: A deep learning system for multi-animal pose tracking. Nature Methods, 19:486-495.

R Core Team. (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/.

West-Eberhard MJ. (1969) The social biology of Polistine wsaps. Miscellaneous Publications, Museum of Zoology, University of Michigan. 15:1-101.

Lorenzi MC, Bagnères AG, Clément JL, Turillazzi S. (1997) *Polistes biglumis bimaculatus* epicuticular hydrocarbons and nestmate recognition (Hymenoptera, Vespidae). Insectes sociaux 44:123-138.

Tumulty JP, Miller SE, Van Belleghem SM, Weller HI, Jernigan CM, Vincent S, Staudenraus RJ, Legan AW, Polnaszek TJ, Uy FMK, Walton A, Sheehan MJ. (2021) Evidence for a selective link between cooperation and individual recognition. bioRxiv DOI: https://doi.org/10.1101/2021.09.07.459327

Dell AI, Bender JA, Branson K, Couzin ID, de Polavieja GG, Noldus LPJJ, Pérez-Escudero A, Perona P, Straw AD, Wikelski M, Brose U. (2014) Automated image-based tracking and its application in ecology. Cell Press 29: 417-428.

Crall JD, Gravish N, Mountcastle AM, Kocher SD, Oppenheimer RL, Pierce NE, Combes SA. (2018) Spatial fidelity of workers predicts collective response to disturbance in a social insect. Nature Communications 9:1201.

Mathis MW, Mathis A. (2020) Deep learning tools for the measurement of animal behavior in neuroscience. Current Opinion in Neurobiology 60:1-11.

Wang ZY, McKenzie-Smith GC, Liu W, Cho HJ, Pereira T, Dhanerawala Z, Shaevitz JW, Kocher SD. (2022) Isolation disrupts social interactions and destabilizes brain development in bumblebees. Current Biology 32:1-11.

Dahake A, Stöckl AL, Foster JJ, Sane SP, Kelber A. (2018) The roles of vision and antennal mechanoreception in hawkmoth flight control. eLife 7:e37606.

Deora T, Ahmed MA, Daniel TL, Brunton BW. (2021) Tactile active sensing in an insect plant pollinator. Journal of Experimental Biology 224(4):jeb239442.

Altshuler DL, Dickson WB, Vance JT, Roberts SP, Dickinson MH. (2005) Short-amplitude high-frequency wing strokes determine the aerodynamics of honeybee flight. Proceedings of the National Academy of Sciences USA 102(50):18213-18218.

Muijres FT, Elzinga MJ, Melis JM, Dickinson MH. (2014) Flies evade looming targets by executing rapid visually directed banked turns. Science 344(6180):172-177.

Peters JM, Gravish N, Combes SA. (2017) Wings as impellers: honey bees co-opt flight system to induce nest ventilation and disperse pheromones. Journal of Experimental Biology 220(12):2203-2209.

Marshall JD, Aldarondo DE, Dunn TW, Wang WL, Berman GJ, Ölveczky BP. (2021) Continuous whole-body 3D kinematic recordings across the rodent behavioral repertoire. Neuron 109(3):420-437.

Ebbesen CL, Froemke RC. (2022) Automatic mapping of multiplexed social receptive fields by deep learning and GPU-accelerated 3D videography. Nature Communications 13:593.

Nunes TM, Nascimento FS, Turatti IC, Lopes NP, Zucchi R. (2008) Nestmate recognition in a stingless bee: does the similarity of chemical cues determine guard acceptance? Animal Behaviour 75(3):1165-1171.

Van Zweden JS, d'Ettorre P. (2010) Nestmate recognition in social insects and the role of hydrocarbons. In Blomquist GJ, Bagnères A-G, eds. Insect Hydrocarbons: Biology, Biochemistry, and Chemical Ecology. Cambridge University Press.

Tibbetts EA. (2002) Visual signals of individual identity in the wasp *Polistes fuscatus*. Proceedings of the Royal Society B 269:1423-1428.

Ortiz CC, Tibbetts EA. 2020. Visual and chemical signals provide different information in *Polistes fuscatus* wasps. Ethology. 127:231-237.

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

Dani FR, Jones GR, Destri S, Spencer SH, Turillazzi S. (2001) Deciphering the recognition signature within the cuticular chemical profile of paper wasps. Animal Behaviour 62(1):165-171.

Bruschini C, Cervo R, Cini A, Pieraccini G, Pontieri L, Signorotti L, Turillazzi S. (2011) Cuticular hydrocarbons rather than peptides are responsible for nestmate recognition in *Polistes dominulus*. Chemical Senses 36(8):715-723.

Nascimento DL, Nascimento FS. (2012) Acceptance threshold hypothesis is supported by chemical similarity of cuticular hydrocarbons in a stingless bee, *Melipona asilvai*. Journal of Chemical Ecology 28:1432-1440.

Cini A, Cappa F, Pepiciello I, Platania L, Dapporto L, Cervo R. (2019) Sight in a clique, scent in a society: plasticity in the use of nestmate recognition cues along colony development in the social wasp *Polistes dominula*. Frontiers in Ecology and Evolution 7:444.

Zhou C, Rao Y, Rao Y. (2008) A subset of octopaminergic neurons are important for *Drosophila* aggression. Nature Neuroscience 11:1059-1067.

Hoopfer ED, Jung Y, Inagaki HK, Rubin GM, Anderson DJ. (2015) P1 interneurons promote a persistent internal state that enhances inter-male aggression in *Drosophila*. eLife 4:e11346.

Anderson DJ, Adolphs R. (2014) A framework for studying emotions across species. Cell 157(1):187-200.

Scharf HM, Suarez AV, Reeve HK, Hauber ME. (2020) The evolution of conspecific acceptance threshold models. Philosophical Transactions of the Royal Society B. 375(1802):20190475.

Rossi N, Baracchi D, Giurfa M, d'Ettorre P. (2019) Pheromone-induced accuracy of nestmate recognition in carpenter ants: simultaneous decrease in type I and type II errors. The American Naturalist 193:267-278.

APPENDIX

Table S1.1. List of peaks detected in *P. fuscatus* using SPME

		2019					2020				
# [a]	Compound(s)	RI	Mean R.A.[b]	SD	Freq. [c]	ANOVA P-value[d]	RI	Mean R.A.	SD	Freq.	ANOVA P-value
1	C21*	2101	0.56	0.93	96	0.33960	2091	0.22	0.81	59	0.44960
2	UNK_01	2142	0.11	0.16	99	0.07690	2136	0.14	0.15	100	0.00693
3	UNK_02	2161	0.73	0.92	100	0.20830	2158	0.22	0.25	100	0.07181
4	UNK_03	2183	0.03	0.04	57	0.80090	2183	0.04	0.06	60	0.00570
5	C22*	2201	0.27	0.35	100	0.07681	2201	0.22	0.26	99	0.00529
6	UNK_04	2279	0.34	0.61	79	0.43990	2278	0.55	0.87	74	0.00028
7	C23*	2302	2.68	2.84	100	0.00802	2301	4.02	3.54	100	0.00029
8	9-MeC23 11-MeC23	2339	0.08	0.06	86	0.00049	2337	0.12	0.10	100	0.00059
9	3-MeC23	2374	0.08	0.07	89	0.00839	2373	0.08	0.06	90	0.00000
10	C24*	2401	0.17	0.10	100	0.00002	2400	0.20	0.11	97	0.00000
11	4-MeC24	2455	NA	NA	6	NA	2460	NA	NA	38	NA
12	C25:1 2-MeC24	2475	0.13	0.23	75	0.31130	2473	0.18	0.21	88	0.00168
13	C25:1	2485	0.08	0.10	61	0.00731	2483	0.15	0.24	77	0.00975
14	C25	2504	7.00	3.45	94	0.00182	2503	9.11	4.22	99	0.00172
15	9-MeC25 11-MeC25 13-MeC25	2537	0.14	0.07	97	0.00034	2535	0.18	0.09	99	0.00001
16	5-MeC25	2558	0.69	1.22	84	0.12670	2554	0.52	0.76	100	0.02738
17	9,13-diMeC25	2568	NA	NA	15	NA	2564	NA	NA	4%	NA
18	3-MeC25	2575	0.10	0.11	86	0.13520	2573	0.12	0.11	100	0.00062
19	C26*	2602	0.36	0.16	100	0.00003	2600	0.47	0.28	98	0.00172
20	9-MeC26 10-MeC26	2630	NA	NA	6	NA	2634	NA	NA	24	NA
21	6-MeC26	2648	NA	NA	11	NA	2649	NA	NA	8	NA
22	4-MeC26	2659	NA	NA	18	NA	2655	NA	NA	16	NA
23	X,Y-diMeC26	NA	NA	NA	0	NA	2668	NA	NA	23	NA
24	3-MeC26	2671	NA	NA	29	NA	2673	NA	NA	38	NA
25	C27:1	2679	NA	NA	15	NA	2680	NA	NA	25	NA
26	C27:1	2692	0.31	1.32	52	0.05248	2690	0.07	0.08	66	0.00055
27	C27*	2703	4.47	3.11	94	0.00001	2702	7.20	4.47	98	0.02414
28	UNK_05	2718	0.23	1.92	56	0.10430	2716	NA	NA	48	NA

29	9-MeC27 11-MeC27 13-MeC27	2735	0.24	0.10	100	0.00481	2735	0.24	0.11	98	0.00550
30	7-MeC27	NA	NA	NA	0	NA	2742	NA	NA	3	NA
31	5-MeC27	2753	0.04	0.04	71	0.14960	2751	0.05	0.03	85	0.00067
32	9,13-diMeC27	2763	0.04	0.05	61	0.58560	2761	0.04	0.05	59	0.00000
33	3-MeC27	2776	0.06	0.05	82	0.21820	2774	0.07	0.09	84	0.31520
34	5,11-diMeC27	2785	NA	NA	32	NA	2782	NA	NA	29	NA
35	5,19-diMeC27	2794	NA	NA	37	NA	2791	NA	NA	32	NA
36	C28	2802	0.17	0.14	87	0.00000	2800	0.32	0.25	99	0.00166
37	5,11,19- triMeC27	NA	NA	NA	0	NA	2817	NA	NA	6	NA
38	9-MeC28 10-MeC28 11-MeC28 12-MeC28 13-MeC28 14-MeC28	2834	0.10	0.27	92	0.34980	2832	0.06	0.04	86	0.00185
39	6-MeC28	NA	NA	NA	0	NA	2848	NA	NA	3	NA
40	4-MeC28	2861	0.03	0.03	65	0.43960	2859	0.06	0.09	87	0.00268
41	3-MeC28	2878	0.03	0.02	72	0.00491	2876	0.07	0.14	82	0.62210
42	C29:1*	2893	0.48	0.55	100	0.00044	2891	0.61	0.45	98	0.00011
43	C29*	2901	1.67	1.54	100	0.00193	2901	3.13	2.63	100	0.00126
44	X-MeC29:1	2924	0.10	0.17	67	0.02372	2922	0.19	0.15	94	0.00133
45	9-MeC29 11-MeC29 13-MeC29 15-MeC29	2935	3.14	1.53	89	0.51310	2933	2.96	1.20	100	0.00000
46	7-MeC29	NA	NA	NA	0	NA	2949	NA	NA	4	NA
47	11,15- diMeC29*	2962	3.06	1.46	100	0.00000	2959	2.16	1.51	100	0.00000
48	9,17-diMeC29	NA	NA	NA	0	NA	2969	0.20	0.16	71	0.00000
49	3-MeC29	2972	0.19	0.21	53	0.00105	2981	0.11	0.14	82	0.78590
50	5,17-diMeC29 5,19-diMeC29 5,21-diMeC29	2984	0.13	0.15	89	0.44260	2985	NA	NA	2	NA
51	C30*	2996	0.15	0.12	98	0.59240	2996	0.24	0.22	100	0.00012
52	3,11-diMeC29 3,13-diMeC29 3,15-diMeC29 3,17-diMeC29 3,19-diMeC29 3,21-diMeC29	3005	0.20	0.12	95	0.01013	3003	0.13	0.11	71	0.13470

53	9-MeC30 10-MeC30 11-MeC30 12-MeC30 13-MeC30 14-MeC30 15-MeC30	3031	0.50	0.39	89	0.11960	3030	0.57	0.21	100	0.00111
54	3,7,Y-triMeC29	3036	0.46	0.42	66	0.05841	3035	NA	NA	33	NA
55	3,7,19-triMeC29	3038	NA	NA	7	NA	NA	NA	NA	0	NA
56	11,Y-diMeC30 12,Z-diMeC30 13,X-diMeC30	3059	0.88	0.33	100	0.00000	3057	0.60	0.26	99	0.00000
57	8,Z-diMeC30 11,Z-diMeC30	3081	0.26	0.15	98	0.00000	3077	0.35	0.27	98	0.00437
58	4,10-diMeC30 4,16-diMeC30	3091	0.23	0.26	85	0.01647	3091	0.40	0.26	95	0.00164
59	C31*	3100	0.47	0.58	78	0.00011	3099	1.40	1.19	97	0.00002
60	X-MeC31:1	3118	NA	NA	30	NA	3113	NA	NA	31	NA
61	Y-MeC31:1	3126	2.13	4.34	63	0.09439	3123	1.62	1.25	89	0.24630
62	9-MeC31 11-MeC31 13-MeC31* 15-MeC31*	3136	10.9	4.06	92	0.36220	3132	10.79	2.35	99	0.00062
63	11,15-diMeC31* 13,17-diMeC31*	3164	19.2	7.70	92	0.00041	3160	16.18	5.32	100	0.00000
64	5,13-diMeC31 5,15-diMeC31 5,17-diMeC31 5,19-diMeC31 5,21-diMeC31 5,23-diMeC31	3186*	3.89	5.72	100	0.00983	3182	2.11	1.22	100	0.00000
65	9,15,19-triMeC31	3196	NA	NA	35	NA	3192	NA	NA	47%	NA
66	3,15-diMeC31 3,17-diMeC31 3,19-diMeC31 3,21-diMeC31 5,9,21-triMeC31	3208	1.62	0.75	96	0.00016	3205*	1.34	0.61	100	0.00000
67	UNK_06	3234	1.93	1.02	93	0.00012	3231	1.77	0.88	100	0.00094
68	3,7,11,15-tetraMeC31 13,Y-diMeC32 14,Z-diMeC32	3260	1.66	0.58	100	0.00000	3257	1.35	0.49	99	0.00012
69	3,7,11,19-tetraMeC31	3270	NA	NA	6	NA	NA	NA	NA	0	NA
70	UNK_07	3280	0.47	0.23	96	0.00000	3277	0.56	0.42	100	0.01179

71	UNK_08	3296	0.08	0.10	69	0.00191	3293	0.10	0.08	80	0.18460
72	6,12,18-triMeC32	3301	0.08	0.10	59	0.02031	3299	0.19	0.20	83	0.00001
73	UNK_09	3308	0.17	0.20	86	0.26460	3305	0.18	0.18	68	0.17460
74	4,8,Y-triMeC32	3326	NA	NA	15	NA	NA	NA	NA	0	NA
75	9-MeC33 11-MeC33 13-MeC33* 15-MeC33* 17-MeC33*	3333	3.84	1.46	91	0.32360	3329	4.47	0.88	100	0.00000
76	11,Y-diMeC33 13,17-diMeC33* 15,19-diMeC33	3362	10.8	4.94	89	0.00432	3359	10.95	3.17	100	0.00005
77	9,21-diMeC33	3370	NA	NA	15	NA	NA	NA	NA	0	NA
78	9,13,17-triMeC33	3381*	1.75	1.09	99	0.00000	3378	1.22	0.65	100	0.00000
79	7,11,19-triMeC33	3393	0.23	0.25	52	0.00000	3391	0.24	0.24	64	0.18700
80	5,11,19-triMeC33 3,21-diMeC33 3,19-diMeC33	3405	0.55	0.28	96	0.00000	3403	0.38	0.22	89	0.00073
81	13-MeC34 15-MeC34 17-MeC34 3,9,19-triMeC33	3433	0.98	0.47	100	0.00006	3430	0.70	0.39	99	0.00000
82	3,7,11,19-tetraMeC33	3458	0.64	0.36	99	0.00000	3455*	0.47	0.33	98	0.00001
83	UNK_10	3481	0.03	0.03	73	0.00186	3478	0.05	0.05	71	0.00003
84	UNK_11	NA	NA	NA	0	NA	3499	0.03	0.07	52	0.39840
85	UNK_12	3507	0.05	0.12	62	0.04697	3505	NA	NA	47	NA
86	4,X,Y-triMeC34	3515	NA	NA	31	NA	3513	NA	NA	18	NA
87	11-MeC35 13-MeC35 15-MeC35 17-MeC35	3528*	0.94	0.64	96	0.00020	3525*	0.96	0.27	100	0.00000
88	11,Y-diMeC35 13,21-diMeC35 15,19-diMeC35	3558	2.54	0.67	100	0.00000	3554	2.34	0.45	100	0.18370
89	11,Z-diMeC35 13,Y-diMeC35	3585	0.28	0.25	79	0.00002	3580	0.38	0.15	99	0.85240
90	5,11,19-triMeC35	3604	0.32	0.30	95	0.00029	3600	0.15	0.11	90	0.08464
91	3,11,Y-triMeC35	3631	0.28	0.11	99	0.00069	3627	0.27	0.16	96	0.00000
92	3,7,11,19-tetraMeC35	3655	0.33	0.21	85	0.03613	3654	0.27	0.24	65	0.23110

93	13,Y-diMeC37 15,7-diMeC37	3680	NA	NA	24	NA	3679	NA	NA	23	NA
94	UNK_13	3725	0.29	0.13	96	0.00000	3722	0.36	0.16	100	0.00000
95	UNK_14	3754	2.54	0.79	100	0.57760	3753	2.66	1.12	100	0.00677
96	UNK_15	3779	0.23	0.16	87	0.00029	3777	0.11	0.15	51	0.00020

a. Peak number
b. Mean relative abundance (as percentage) of the peak in the peak table of peaks present in >50% of individuals
c. Frequency (as a percentage) of the peak among wasps sampled each year
d. Underlined P-values are significant after Bonferroni correction for multiple comparisons.

- Asterisks in the Compound(s) column indicate this compound was found to be diagnostic of nest affiliation by either Espelie et al. (1994) or Gamboa et al. (1996).
- Underlined retention indices (RI) designate the peaks in the top 3 canonical loadings of discriminant functions 1 and 2 in DAPC analyses.

Asterisks on RI indicate the top 3 peaks diagnostic of nest in random forest analyses.

Table S3.1. Distributions of OR genes across WGCNA modules

gene	Module Color	GS.Male	p.GS.Male	GS.Female	p.GS.Female	OR_subfam
PfusOR4	black	-0.0464137	0.89221024	0.04641374	0.89221024	H
PfusOR175	black	-0.1280154	0.70758209	0.12801543	0.70758209	9-exon
PfusOR28	black	-0.4591139	0.15545727	0.45911395	0.15545727	9-exon
PfusOR15	black	-0.1478514	0.66440711	0.14785143	0.66440711	F
PfusOR137	black	-0.3032786	0.36461659	0.30327864	0.36461659	9-exon
PfusOR106	black	0.00699755	0.98370918	-0.0069975	0.98370918	9-exon
PfusOR167	black	0.02847501	0.93376688	-0.028475	0.93376688	9-exon
PfusOR127	black	0.11493138	0.7364987	-0.1149314	0.7364987	9-exon
PfusOR102PSE	black	-0.1112561	0.74467844	0.11125614	0.74467844	OR
PfusOR40	black	0.04542375	0.89449819	-0.0454237	0.89449819	I
PfusOR72	black	0.26722102	0.42697796	-0.267221	0.42697796	D
PfusOR103PSE	black	-0.0111388	0.9740703	0.0111388	0.9740703	9-exon
PfusOR135	black	-0.0312371	0.9273562	0.03123711	0.9273562	9-exon
PfusOR163	black	-0.5328193	0.09149415	0.53281926	0.09149415	E
PfusOR45	blue	0.58756323	0.05732652	-0.5875632	0.05732652	H

PfusOR 172	blue	0.55213458	0.07821386	-0.5521346	0.07821386	9-exon
PfusOR 66	blue	-0.4571911	0.15741997	0.45719107	0.15741997	T
PfusOR 82	blue	-0.4493793	0.16555132	0.44937935	0.16555132	9-exon
PfusOR 202	blue	0.47735324	0.13759824	-0.4773532	0.13759824	L
PfusOR 52NTE	blue	-0.4986509	0.11845796	0.49865093	0.11845796	9-exon
PfusOR 55	blue	-0.6360795	0.0353946	0.63607954	0.0353946	T
PfusOR 159	blue	0.33526603	0.31350541	-0.335266	0.31350541	E
PfusOR 197	blue	0.46327541	0.15126202	-0.4632754	0.15126202	L
PfusOR 199	blue	0.45226209	0.16252109	-0.4522621	0.16252109	L
PfusOR 32	blue	-0.5763949	0.06344217	0.57639492	0.06344217	9-exon
PfusOR 198	blue	0.43991008	0.17574906	-0.4399101	0.17574906	L
PfusOR 22NC	blue	0.4264813	0.190855	-0.4264813	0.190855	9-exon
PfusOR 124	blue	-0.5180625	0.10258458	0.51806248	0.10258458	9-exon
PfusOR 189	blue	0.36095917	0.2754466	-0.3609592	0.2754466	L
PfusOR 222	blue	0.66677381	0.02503905	-0.6667738	0.02503905	9-exon
PfusOR 56	blue	-0.4154613	0.20381795	0.41546131	0.20381795	T
PfusOR 190	blue	0.74990812	0.00786425	-0.7499081	0.00786425	L
PfusOR 54	blue	-0.2040259	0.54734424	0.20402589	0.54734424	T
PfusOR 195	blue	0.76038962	0.00659402	-0.7603896	0.00659402	L
PfusOR 177	blue	-0.4735881	0.14117302	0.47358815	0.14117302	K
PfusOR 68PSE	brown	-0.2320155	0.49240089	0.23201548	0.49240089	OR
PfusOR 46	brown	0.13749397	0.68684606	-0.137494	0.68684606	H
PfusOR 14	brown	0.10585644	0.75673904	-0.1058564	0.75673904	F
PfusOR 27	brown	0.10080678	0.7680621	-0.1008068	0.7680621	9-exon
PfusOR 114	brown	0.09053242	0.79122469	-0.0905324	0.79122469	9-exon
PfusOR 208	brown	0.08724185	0.79867604	-0.0872419	0.79867604	V
PfusOR 193	brown	0.15660982	0.64562269	-0.1566098	0.64562269	L

PfusOR 164	brown	-0.2290234	0.49815527	0.22902338	0.49815527	E
PfusOR 96	brown	0.11195433	0.74312266	-0.1119543	0.74312266	J
PfusOR 93	brown	0.5430562	0.08428393	-0.5430562	0.08428393	9-exon
PfusOR 122	cyan	-0.3680953	0.2653596	0.36809533	0.2653596	9-exon
PfusOR 105	cyan	-0.3084331	0.35610702	0.30843307	0.35610702	9-exon
PfusOR 117	cyan	-0.1269213	0.70998741	0.12692134	0.70998741	9-exon
PfusOR 74	darkgreen	-0.0911347	0.7898625	0.09113472	0.7898625	9-exon
PfusOR 129	darkgreen	-0.1530352	0.65326752	0.15303523	0.65326752	9-exon
PfusOR 149	darkgrey	0.14620924	0.66794891	-0.1462092	0.66794891	9-exon
PfusOR 13	darkgrey	0.28554756	0.39467156	-0.2855476	0.39467156	F
PfusOR 48	darkgrey	0.50173707	0.1158352	-0.5017371	0.1158352	H
PfusOR 151	darkgrey	0.20672197	0.54194602	-0.206722	0.54194602	9-exon
PfusOR 182	darkgrey	0.35631389	0.28212632	-0.3563139	0.28212632	L
PfusOR 192	darkgrey	0.74730652	0.00820539	-0.7473065	0.00820539	L
PfusOR 211	darkgrey	0.4943422	0.12218314	-0.4943422	0.12218314	V
PfusOR 145PSE	darkmagenta	-0.6581329	0.02770236	0.65813293	0.02770236	OR
PfusOR 29	darkolivegreen	0.13173562	0.69942138	-0.1317356	0.69942138	W
PfusOR 95	darkolivegreen	-0.3094438	0.3544505	0.30944383	0.3544505	J
PfusOR 50PSE	darkorange	0.11019087	0.74705381	-0.1101909	0.74705381	H
PfusOR 201	darkred	0.11590893	0.73432711	-0.1159089	0.73432711	L
PfusOR 156	darkred	-0.160149	0.63808379	0.160149	0.63808379	9-exon
PfusOR 47	darkred	0.41178469	0.20825668	-0.4117847	0.20825668	H
PfusOR 11	darkturquoise	0.2869395	0.39226887	-0.2869395	0.39226887	F
PfusOR 203	green	-0.4605703	0.15398094	0.46057031	0.15398094	R
PfusOR 200	green	0.0274779	0.93608202	-0.0274779	0.93608202	L
PfusOR 26	green	0.48232072	0.1329702	-0.4823207	0.1329702	9-exon
PfusOR 69	green	-0.4803271	0.13481551	0.48032713	0.13481551	Q

PfusOR12	green	-0.3038857	0.36360896	0.30388573	0.36360896	F
PfusOR57	green	-0.0481104	0.88829093	0.04811037	0.88829093	T
PfusOR53	green	-0.4442407	0.17103881	0.44424068	0.17103881	T
PfusOR58	green	-0.5946994	0.05363659	0.59469939	0.05363659	T
PfusOR65	green	-0.3440349	0.30021181	0.34403494	0.30021181	T
PfusOR108	grey	-0.3790406	0.25030118	0.37904059	0.25030118	9-exon
PfusOR78PSE	grey60	-0.1686967	0.6200036	0.16869669	0.6200036	9-exon
PfusOR30PSE	lightcyan	0.0814951	0.81172518	-0.0814951	0.81172518	OR
PfusOR70	lightcyan	0.10552435	0.75748239	-0.1055244	0.75748239	9-exon
PfusOR64	lightgreen	0.19126481	0.57318608	-0.1912648	0.57318608	T
PfusOR170	magenta	-0.2253597	0.50524093	0.22535972	0.50524093	9-exon
PfusOR49	magenta	0.32220677	0.33388019	-0.3222068	0.33388019	H
PfusOR224	magenta	0.04977815	0.88444044	-0.0497782	0.88444044	A
PfusOR216	magenta	0.33708793	0.31071766	-0.3370879	0.31071766	9-exon
PfusOR94	magenta	0.79425793	0.0035017	-0.7942579	0.0035017	9-exon
PfusOR31	midnightblue	-0.392431	0.23256179	0.39243104	0.23256179	9-exon
PfusOR79	midnightblue	-0.4933977	0.12300963	0.4933977	0.12300963	9-exon
PfusOR10	midnightblue	-0.2375528	0.48182957	0.23755277	0.48182957	P
PfusOR75	midnightblue	-0.4978813	0.1191179	0.49788134	0.1191179	9-exon
PfusOR154	midnightblue	-0.4942722	0.12224425	0.49427223	0.12224425	9-exon
PfusOR174	midnightblue	-0.3950562	0.22917244	0.39505621	0.22917244	9-exon
PfusOR100PSE	midnightblue	-0.6442409	0.03239332	0.64424094	0.03239332	9-exon
PfusOR128	midnightblue	-0.8168759	0.00214702	0.81687586	0.00214702	9-exon
PfusOR152	midnightblue	0.14328517	0.67427046	-0.1432852	0.67427046	9-exon
PfusOR125	midnightblue	-0.3255257	0.32863706	0.32552572	0.32863706	9-exon
PfusOR229CTE	orange	0.23824072	0.48052332	-0.2382407	0.48052332	S
PfusOR35	orange	0.14124575	0.67869071	-0.1412458	0.67869071	9-exon

PfusOR158	orange	0.1880008	0.57987113	-0.1880008	0.57987113	E
PfusOR38	orange	-0.0368931	0.91424149	0.03689306	0.91424149	Unclassified
PfusOR196	orangered4	-0.061214	0.85810222	0.06121403	0.85810222	L
PfusOR17	paleturquoise	-0.2935106	0.38102517	0.29351061	0.38102517	Z
PfusOR138PN	paleturquoise	-0.7416638	0.00898226	0.74166384	0.00898226	9-exon
PfusOR123	paleturquoise	-0.4961091	0.12064662	0.49610906	0.12064662	9-exon
PfusOR41	paleturquoise	-0.5201351	0.10097658	0.52013508	0.10097658	9-exon
PfusOR42	pink	0.27977512	0.40471323	-0.2797751	0.40471323	9-exon
PfusOR3	pink	0.5208867	0.10039754	-0.5208867	0.10039754	G
PfusOR185	pink	-0.857452	0.00073988	0.85745203	0.00073988	L
PfusOR144	plum1	-0.5694352	0.06746837	0.5694352	0.06746837	9-exon
PfusOR111	purple	-0.0001833	0.99957313	0.00018335	0.99957313	9-exon
PfusOR59	purple	0.17874862	0.59898131	-0.1787486	0.59898131	T
PfusOR217	purple	0.31577725	0.3441622	-0.3157773	0.3441622	9-exon
PfusOR213	purple	0.43373439	0.18260202	-0.4337344	0.18260202	9-exon
PfusOR221	red	0.58194365	0.06035115	-0.5819437	0.06035115	9-exon
PfusOR60	red	0.16948845	0.61833818	-0.1694885	0.61833818	T
PfusOR178	red	0.33411593	0.31527214	-0.3341159	0.31527214	L
PfusOR204	red	0.72344335	0.01185799	-0.7234434	0.01185799	C
PfusOR168	royalblue	-0.1570671	0.644647	0.15706707	0.644647	9-exon
PfusOR210	royalblue	0.27469762	0.41364848	-0.2746976	0.41364848	V
PfusOR8	royalblue	-0.5979619	0.05200508	0.59796185	0.05200508	O
PfusOR73	royalblue	-0.2324636	0.49154163	0.23246358	0.49154163	9-exon
PfusOR206	royalblue	0.27225649	0.41797813	-0.2722565	0.41797813	Q
PfusOR97	saddlebrown	0.46903055	0.14557785	-0.4690306	0.14557785	J
PfusOR21	saddlebrown	0.32283876	0.3328784	-0.3228388	0.3328784	9-exon
PfusOR37	salmon	-0.2357372	0.4852845	0.23573719	0.4852845	9-exon

PfusOR209NTE	salmon	0.34688905	0.29595274	-0.3468891	0.29595274	V
PfusOR228PSE	salmon	-0.3427672	0.30211432	0.34276719	0.30211432	OR
PfusOR112	salmon	-0.2126751	0.53010451	0.21267508	0.53010451	9-exon
PfusOR118	salmon	-0.673279	0.02315747	0.67327904	0.02315747	9-exon
PfusOR121PN	salmon	-0.4759178	0.13895426	0.47591784	0.13895426	9-exon
PfusOR132	salmon	-0.2630902	0.43442947	0.2630902	0.43442947	9-exon
PfusOR104	salmon	-0.4598558	0.15470415	0.4598558	0.15470415	9-exon
PfusOR139PSE	salmon	-0.3648087	0.26997901	0.36480865	0.26997901	OR
PfusOR176	salmon	-0.1808938	0.59452958	0.18089383	0.59452958	9-exon
PfusOR142PSE	salmon	-0.3842282	0.24333941	0.38422816	0.24333941	OR
PfusOR5	salmon	-0.2482067	0.46178027	0.24820667	0.46178027	H
PfusOR140	salmon	-0.1413531	0.67845779	0.14135311	0.67845779	9-exon
PfusOR133	salmon	-0.3113628	0.35131657	0.31136279	0.35131657	9-exon
PfusOR7NTE	salmon	-0.180213	0.59594108	0.180213	0.59594108	O
PfusOR162	skyblue	-0.2582015	0.44332708	0.2582015	0.44332708	E
PfusOR92	skyblue	0.47080899	0.14384889	-0.470809	0.14384889	9-exon
PfusOR115	steelblue	-0.0352958	0.91794333	0.03529579	0.91794333	9-exon
PfusOR148	turquoise	0.14026807	0.68081299	-0.1402681	0.68081299	9-exon
PfusOR166	turquoise	0.09120972	0.78969291	-0.0912097	0.78969291	9-exon
PfusOR131	turquoise	0.06077771	0.85910489	-0.0607777	0.85910489	9-exon
PfusOR146	turquoise	0.00956581	0.97773117	-0.0095658	0.97773117	9-exon
PfusOR76	turquoise	0.06737662	0.8439611	-0.0673766	0.8439611	9-exon
PfusOR101	turquoise	-0.1611922	0.63586749	0.16119218	0.63586749	9-exon
PfusOR77	turquoise	0.12698908	0.7098384	-0.1269891	0.7098384	9-exon
PfusOR67	turquoise	-0.3456162	0.29784798	0.34561624	0.29784798	9-exon
PfusOR160	turquoise	0.40853065	0.21223277	-0.4085307	0.21223277	E
PfusOR33	turquoise	-0.0172467	0.95985989	0.01724675	0.95985989	9-exon

PfusOR34	turquoise	0.05576171	0.87064493	-0.0557617	0.87064493	9-exon
PfusOR165	turquoise	0.18740815	0.58108815	-0.1874081	0.58108815	9-exon
PfusOR173	turquoise	0.47736796	0.13758437	-0.477368	0.13758437	9-exon
PfusOR130	turquoise	0.15166765	0.65620029	-0.1516676	0.65620029	9-exon
PfusOR150	turquoise	0.05434097	0.87391775	-0.054341	0.87391775	9-exon
PfusOR16	turquoise	0.42469934	0.19291649	-0.4246993	0.19291649	9-exon
PfusOR169	turquoise	0.266871	0.42760697	-0.266871	0.42760697	9-exon
PfusOR141	turquoise	-0.1142547	0.73800289	0.11425471	0.73800289	9-exon
PfusOR61	turquoise	0.37508684	0.25568292	-0.3750868	0.25568292	T
PfusOR36	turquoise	0.49813197	0.11890272	-0.498132	0.11890272	9-exon
PfusOR113PSE	turquoise	0.49146571	0.12471136	-0.4914657	0.12471136	OR
PfusOR91	turquoise	-0.3739382	0.25725865	0.37393822	0.25725865	9-exon
PfusOR205	turquoise	0.3146157	0.34603723	-0.3146157	0.34603723	Q
PfusOR9	turquoise	0.04145414	0.90367934	-0.0414541	0.90367934	O
PfusOR161	turquoise	-0.1733046	0.61033389	0.1733046	0.61033389	E
PfusOR191	turquoise	0.1089497	0.74982389	-0.1089497	0.74982389	L
PfusOR207	turquoise	0.44576193	0.16940281	-0.4457619	0.16940281	V
PfusOR116	turquoise	0.19840806	0.55866207	-0.1984081	0.55866207	9-exon
PfusOR223	turquoise	-0.2355555	0.48563075	0.23555555	0.48563075	9-exon
PfusOR171	turquoise	0.27486335	0.41335533	-0.2748633	0.41335533	9-exon
PfusOR184	turquoise	0.48542326	0.13013046	-0.4854233	0.13013046	L
PfusOR126	turquoise	-0.1865027	0.58294933	0.18650274	0.58294933	9-exon
PfusOR186	turquoise	0.64137576	0.03342581	-0.6413758	0.03342581	L
PfusOR214PSE	turquoise	0.56603176	0.06949851	-0.5660318	0.06949851	OR
PfusOR215	turquoise	0.15758552	0.64354129	-0.1575855	0.64354129	9-exon
PfusOR183NTE	turquoise	0.49682216	0.12003002	-0.4968222	0.12003002	L
PfusOR181	turquoise	0.42927263	0.18765269	-0.4292726	0.18765269	L

PfusOR218	turquoise	-0.0530986	0.87678117	0.05309858	0.87678117	9-exon
PfusOR179	turquoise	0.50552681	0.11266606	-0.5055268	0.11266606	L
PfusOR180	turquoise	0.5134591	0.10621552	-0.5134591	0.10621552	L
PfusOR6	violet	0.32681283	0.32661562	-0.3268128	0.32661562	M
PfusOR187	violet	-0.3526764	0.28741916	0.35267644	0.28741916	L
PfusOR86	white	0.28326527	0.39862691	-0.2832653	0.39862691	9-exon
PfusOR157	white	0.30874777	0.35559084	-0.3087478	0.35559084	E
PfusOR119	yellow	-0.4647419	0.1498007	0.46474185	0.1498007	9-exon
PfusOR143	yellow	-0.7289405	0.01092888	0.72894047	0.01092888	9-exon
PfusOR98	yellow	0.82296189	0.00186139	-0.8229619	0.00186139	Q
PfusOR136	yellow	-0.8575392	0.00073794	0.85753921	0.00073794	9-exon
PfusOR153	yellow	-0.6982523	0.01685969	0.69825231	0.01685969	9-exon
PfusOR134	yellow	-0.4692972	0.14531776	0.46929724	0.14531776	9-exon
PfusOR147	yellow	-0.5284493	0.09469191	0.52844927	0.09469191	9-exon
PfusOR212	yellow	0.33047082	0.32090699	-0.3304708	0.32090699	V
PfusOR110	yellow	-0.5004802	0.11689877	0.50048022	0.11689877	9-exon
PfusOR107	yellow	0.60125452	0.05039325	-0.6012545	0.05039325	9-exon
PfusOR19	yellow	0.64371182	0.03258229	-0.6437118	0.03258229	9-exon
PfusOR188	yellow	0.82546918	0.00175238	-0.8254692	0.00175238	L
PfusOR71	yellow	0.69763651	0.0169982	-0.6976365	0.0169982	B
PfusOR2	yellow	0.75617845	0.00708474	-0.7561784	0.00708474	G
PfusOR99	yellow	0.80812288	0.00261358	-0.8081229	0.00261358	J
PfusOR43	Yellow green	0.16247609	0.63314344	-0.1624761	0.63314344	9-exon

Table S4.1. Tracking quality and posture measurements for each trial and interval

Trial	Interval	Treatment	Nest_ID	Head_percent_frames_tracked	Thorax_percent_frames_tracked	Abdomen_percent_frames_tracked	LWing_percent_frames_tracked	RWing_percent_frames_tracked	LAnt_percent_frames_tracked	RAnt_percent_frames_tracked	Distance_Head_Traveled	Distance_Thorax_Traveled	Mean_Wing_degrees	Mean_Distance_Wing_Tips	Mean_Ant_Degrees	Mean_Distance_Antennae_Tips
8565	1	Intruder	12	100	100	100	98	100	52	42	13730	14886	5	42	106	205
8565	3	Intruder	12	99	100	95	85	84	85	85	45709	44694	40	310	105	246
8569	1	Dowel	22	83	86	83	77	82	78	79	22312	20821	25	191	113	282
8569	2	Dowel	22	97	97	96	96	96	92	93	27419	26645	30	244	115	284
8569	3	Dowel	22	88	88	88	87	88	83	81	21724	17653	22	170	119	302
8571	1	Dowel	19	57	57	55	47	53	48	47	19081	16716	17	124	108	247
8571	2	Dowel	19	82	78	73	71	71	78	75	23856	16449	18	134	118	240
8571	3	Dowel	19	100	100	99	99	100	97	97	16086	18834	20	152	116	217
8572	1	Intruder	7	100	98	66	95	95	92	89	24982	16631	9	67	105	242
8572	3	Intruder	7	91	89	76	71	75	70	76	162031	104254	44	314	97	199
8573	1	Intruder	15	100	100	99	81	100	95	92	33705	27853	17	120	104	247
8573	3	Intruder	15	78	85	69	63	69	42	52	90437	78104	41	243	100	166
8574	1	Dowel	60	100	100	100	100	100	100	100	7437	4977	7	44	115	231
8574	2	Dowel	60	53	52	47	50	37	41	46	11682	9452	14	102	94	193
8574	3	Dowel	60	24	9	6	5	5	22	6	11720	6882	23	161	98	171
8601	1	Dowel	12	100	100	96	99	97	97	95	23130	23804	27	184	111	211
8601	2	Dowel	12	100	100	99	99	99	98	95	15964	12925	33	226	118	231
8601	3	Dowel	12	99	99	97	97	97	94	96	18242	17386	20	139	112	229
8604	1	Intruder	22	95	96	94	92	92	76	80	39618	28317	17	115	112	226
8604	3	Intruder	22	99	99	96	92	94	79	81	47008	36419	35	220	99	199
8605	1	Dowel	15	98	99	98	95	96	93	92	17129	12871	8	57	116	261
8605	2	Dowel	15	96	95	90	89	87	88	88	43088	36797	31	201	99	210
8605	3	Dowel	15	99	99	98	89	80	87	95	22119	20913	33	214	104	228
8607	1	Dowel	7	96	95	91	84	87	84	85	30114	20307	9	55	101	197
8607	2	Dowel	7	100	100	98	98	96	93	88	49368	33301	30	189	101	186
8607	3	Dowel	7	99	99	98	95	98	94	95	26508	17358	13	81	103	195
8608	1	Intruder	19	95	94	50	48	50	63	40	34940	25577	18	108	112	163
8608	3	Intruder	19	99	99	96	93	98	80	81	54902	40994	37	214	101	159
8609	1	Intruder	60	99	99	95	94	98	93	93	5859	5844	22	121	112	192
8609	3	Intruder	60	95	95	79	88	87	77	79	44809	27604	33	171	102	150