Short title: Byers & Jacobs – Quantitative insect GC-EAD for plant interaction analysis

Title: Quantitative analysis of gas chromatography-coupled electroantennographic detection (GC-EAD) of plant volatiles by insects

Kelsey J.R.P. Byers[1] & Robert N. Jacobs[2]

[1] Department of Cell and Developmental Biology, John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, United Kingdom. To whom correspondence may be addressed: Kelsey.Byers@jic.ac.uk, ORCID: 0000-0002-0923-8264

[2] 6 Greenways, Norwich, NR4 6HE, United Kingdom. Correspondence: robert.n.jacobs@gmail.com

Manuscript category: Protocol Note

Manuscript received _____; revision accepted _____.

Main text number of words: 4048

Abstract

Premise: Plant-interacting insects receive plant volatile signals through their antennae, and voltage changes across an antenna exposed to volatile stimuli can be measured to determine if insects can perceive them, often by coupling gas chromatography to electroantennographic detection (GC-EAD). Current methods for analysing GC-EAD data rely on *qualitative* observation, rather than using *quantitative* analysis, and are thus prone to bias.

Methods and Results: We developed a novel *quantitative* methodology for analysis of GC-EAD data using a signal processing technique on EAD data to compare responses of hawkmoths and bumble bees to a library of common floral volatiles. Responses varied between species and sexes and these responses were in some cases affected by compound type or modification.

Conclusions: Our novel GC-EAD analysis technique is robust to baseline drift and low signal:noise ratios commonly found in GC-EAD data, and provides a way forward for quantitative studies of plant volatile-mediated plant-insect interactions.

Key words: electroantennography; electrophysiology; first forward difference; floral scent; GC-EAD; herbivory; plant volatiles; pollination

32 **INTRODUCTION**

33 Plants produce a wide array of secondary metabolites – i.e. those that are not immediately tied to
34 growth and reproduction – with over 200,000 isolated and described across the plant kingdom (Kessler
35 and Kalske, 2018). Many of these compounds play direct roles in the interactions of plants with the
36 abiotic and biotic environment, and have been exploited for this purpose by humanity, including the
37 agricultural (e.g. geraniol, a terpenoid volatile, which repels eggplant-infesting pest moths (Ghosh et al.,
38 2023; Byers, 2023)) and flavour and fragrance industries (e.g. vanillin, a volatile product of the
39 phenylpropanoid pathway, which is the dominant compound in natural vanilla extract (Gallage et al.,
40 2014)). Although these compounds can be exploited in such ways, and although chemical structures of
41 numerous plant metabolites have been described, their biological functions remain unknown in many
42 cases (Dixon and Dickinson, 2024). Many of these compounds are volatiles, which are especially well
43 known for influencing plants' interactions with insects by attracting both beneficial and antagonistic
44 species (e.g. pollinators and herbivores (Raguso, 2008)), with significant implications for plant ecology
45 and evolution (Byers, 2021) as well as agriculture (Pickett and Khan, 2016).

46 Identifying which volatile cues emitted by plants are perceived by plant-interacting insects is key to our
47 understanding of these cross-kingdom interactions. In particular, identifying insect-relevant volatiles can
48 facilitate work in plant secondary metabolism, ecology, and evolution. For example, identifying pest-
49 resistant crop varieties by screening for known resistance volatiles is cheaper and easier than
50 performing field trials (e.g. Ghosh et al., 2023), while determining which volatiles are key to the ecology
51 and evolution of plant-pollinator interactions can facilitate genetic mapping of only those volatiles
52 rather than the entire plant headspace (e.g. Byers et al., 2014a, 2014b). In short, identifying insect-
53 relevant volatiles – often only a fraction of a plant's emitted volatiles and individually not always the
54 most abundant volatiles (Clavijo McCormick et al., 2014) – can dramatically reduce the experimental
55 scope and increase the tractability of plant secondary metabolite study.

56 Electroantennography – the detection and measurement of electrical signals from whole insect
57 antennae – is often used to screen plant volatiles for insect perceptibility. The compounds to which the
58 insect antenna "responds" with some voltage change are taken to be physiologically active and may be
59 further tested via behavioural assays with single pure or mixed compounds or with plant lines known to
60 be high or low in the relevant compounds. Although these techniques have been in use for nearly
61 seventy years, being first described in 1957 (Schneider, 1957), the resulting data are always presented
62 qualitatively – "response/no response" (e.g. Peakall et al., 2010) – with no quantitative, repeatable
63 measure used to define antennal "response" versus background activity. Indeed, a comprehensive and
64 well-cited overview of GC-EAD includes no mention of how data analysis is performed or what criteria
65 are used to define an "active" peak (Schiestl and Marion-Poll, 2002). This has hampered the field by
66 limiting reproducibility and cross-study or cross-observer comparisons, introducing error, and reducing
67 ease of application to broader plant volatile studies. This may be a particular issue with many plant
68 volatiles, which – unlike insect pheromones – may not have the same degree of "dedicated hardware"
69 on the insect antenna (Akers and Getz, 1993; Schiestl and Marion-Poll, 2002) and thus may produce
70 weaker signals, with the exception of sexually deceptive insect pollination systems which use insect
71 pheromone compounds to attract pollinators.

72 In this Protocol Note, I present "best practice" statistical approaches incorporating techniques from
73 signal processing to analyse gas chromatography-coupled electroantennographic detection (GC-EAD)
74 data in a quantitative fashion. Methods for correcting for baseline drift (specifically, the use of the
75 antennal signal first forward difference), background antennal noise (specifically, comparison with
76 response to the solvent control), and antennal sensitivity loss (via use of windowing) are presented. A
77 sample data set from six *Manduca sexta* hawkmoths (three males and three females), four *Bombus*
78 *terrestris audax* bumble bees (all workers), and 28 common volatile stimuli of different volatile chemical
79 skeleton types (all present in a variety of plant species) are used to demonstrate the methods and their

application in detail and show that they work for a variety of volatile plant secondary metabolites and two major model pollinating insect species. I show that quantitative analyses of electroantennogram data are indeed possible and quite feasible, pointing the way forward for robust, statistically-based analyses of antennal detection ability in volatile-driven plant-insect interactions.

METHODS AND RESULTS

Methods: Gas chromatography-coupled electroantennographic detection (GC-EAD) setup and recording

Insect sourcing

Adult *Manduca sexta* (L.) (Lepidoptera: Sphingidae) hawkmoth males and females were sourced from a maintained colony reared on artificial diet (Frontiers Scientific Services, Newark, Delaware, USA) at the John Innes Centre Entomology platform facility; adults were not fed but females were allowed to oviposit freely on live *Nicotiana benthamiana* plants and males and females were kept together prior to collection. Worker *Bombus terrestris audax* (Harris) (Hymenoptera: Apidae) were sourced from a single Bumblebee New Season Hive (Biobest, purchased from Agralan Ltd, Wiltshire, UK) which were allowed to freely forage in a custom-built acrylic cage with 25% sucrose and ground pollen *ad libitum* (Sevenhills Wholefoods Organic Raw Bee Pollen, Dinnington, UK).

EAD test mixture preparation and analysis

To assess antennal responses to multiple compounds delivered at approximately equal flame ionization detector (FID) sensitivities, a mixture of 28 authentic reference standards was assembled, representing 29 GC-FID peaks (farnesol is sold as a mixture of isomers and both elute separately on the GC-FID). Standards were obtained from a variety of established sources (see Table S1 for vendor addresses and

102 Table S2 for vendor and product numbers) and diluted to 10µg/µl in hexane (HPLC-grade, ca. 95%
103 purity). The final compound list was selected to include a variety of chemical skeletons (terpenoids,
104 aromatics, fatty acid derived compounds, and one nitrogenous compound) that would separate well on
105 a 10°C per minute GC oven ramp program. Due to differences in FID responsiveness to the individual
106 compounds (i.e. some compounds may produce different peak areas on the FID despite being at equal
107 injected amounts), different volumes of the 10µg/µl dilutions were added to the mixture (see Table S2
108 for final concentrations, which ranged from 93ng/µl for (R)-(+)-limonene to 2004ng/µl (2.004µg/µl) for
109 indole).

110 To determine retention times and mixture peak identities, 3ul of the test mixture was manually injected
111 into an Agilent 7890B GC (Agilent, Stockport, UK) fitted with an Agilent 5977A/B MS detector (GC-MS)
112 and a Phenomenex Zebron ZB-5HT column (30m x 0.25mm x 0.10µm; Phenomenex UK Ltd, Macclesfield,
113 United Kingdom). The GC run parameters were as follows: the column flow was 1.0603 ml/min of
114 helium carrier gas with an oven program of 50°C for 4 minutes followed by a temperature ramp of
115 10°C/minute to 230°C, which was then held for 4 minutes. The inlet was splitless and held at 250°C and
116 the MS source was held at 230°C. Compounds were identified using Agilent Unknowns software (v10.1,
117 Agilent, Stockport, UK) and via comparison to Kovats Retention Indices of the individual compounds run
118 separately using the same GC-MS system, calculated using an alkane ladder (49452-U, Sigma-Aldrich,
119 Watford, UK).

120 *EAD antennal preparation and assembly*

121 *Manduca sexta* male and female adults: single antennae were excised using microdissection scissors
122 from three individuals each of living adult male and female *Manduca sexta*. Due to biosecurity
123 requirements for this non-native potential plant pest species, excised antennae were placed in a 15ml
124 conical tube on wet tissue paper to maintain antennal humidity levels and carried to the

125 electrophysiology setup in a separate building. The fine tip of an antenna (representing ca. 7 segments,
126 cut where the tip abruptly broadens) was removed using a microscalpel under a dissection microscope
127 and the antenna mounted on an antenna electrode fork (Syntech PRG-2) using electrode gel (Signa Gel,
128 Parker Laboratories, Fairfield, New Jersey, USA), which was then plugged in to the electrode mount
129 (Syntech MP-15 with EAG Combi Probe) and moved into position in front of the GC eluent-carrying air
130 stream approximately 1cm from the air stream outlet.

131 *Bombus terrestris audax* workers: individuals were removed from the Entomology platform facility in
132 50ml conical tubes, then kept at 4°C for 30 minutes to enable safe antennal removal. Antennae were
133 then removed using microdissection scissors under red light (with the insect held using a set of forceps)
134 and the tip of the final flagellar segment removed using a microscalpel under a dissection microscope,
135 then treated as above.

136 *GC-EAD setup and recording procedure*

137 The GC-FID setup consisted of an Agilent 8890 GC fitted with a flame ionization detector (FID). The
138 column (a Phenomenex Zebron ZB-5HT column (35m x 0.25mm x 0.10µm), identical to the one used for
139 GC-MS analysis) was fitted to a splitless inlet kept at 250°C. The detector end of the column was fitted to
140 the bottom end of an inert glass Y-splitter (Agilent UltraInert Universal PressFit Y-Splitter , part 5190-
141 6980). The two other arms of the splitter were fitted with deactivated column segments of equal 30cm
142 length (Agilent CP803210, 0.32mm inner diameter), with one end running to the FID (which was kept at
143 300°C) and the other end leaving the GC oven through a port on the side which was fitted with a
144 Syntech EC-03 heated transfer tube kept at 250°C. The GC used hydrogen as a carrier gas, which was
145 generated with a Parker Hydrogen Generator (H2PEM-100, Thames Restek UK Ltd, Saunderton, United
146 Kingdom). Nitrogen gas at a flow of 25ml/min was used as a make-up gas, and the FID was also fed with
147 compressed air at 400ml/min to ensure adequate flame gas supply. The GC run parameters were as

148 follows: the column flow was 4.7 ml/min with an oven program of 50°C for 4 minutes followed by a
149 temperature ramp of 10°C/minute to 230°C, which was then held for 4 minutes. Each manual sample
150 injection consisted of 3ul of the EAD test mixture, for a total of 0.279-6.012µg of compound injected.
151 GC-FID chromatogram waveforms were exported from Agilent OpenLab CDS v2.5 software as CSV files
152 consisting of an FID amplitude value (in pA) for each given sampled timepoint.

153 The EAD recording setup consisted of a Syntech IDAC-2 (Ockenfels Syntech, Buchenbach, Germany)
154 connected to the desktop computer controlling the GC-FID system. A constant air flow (setting 8 on the
155 CS-55) was delivered using a Syntech CS-55 Stimulus Controller throughout the experiment. Prior to and
156 shortly after each GC-FID run, a puff of the EAD test mixture (via a Pasteur pipette fitted with filter
157 paper, with 10µl of the EAD test mixture added onto the paper, also setting 8 on the CS-55) was
158 delivered using the CS-55 to ensure antennal viability had been maintained throughout the GC-FID run.
159 EAD signals were captured using Syntech GcEAD 2014 v1.2.5 software with default settings and
160 exported as CSV waveforms in GcEAD 2014 (measured as millivolts) prior to analysis.

161 **Methods: Data processing and analysis**

162 *GC-FID peak identification*

163 A custom Perl script (Perl, 2023; see Appendix for script) was developed to determine the "onset" and
164 "offset" timings (i.e. when a peak began and when it ended) of GC-FID peaks above a certain user-
165 defined threshold from the FID CSV data. In the presented experimental data, the threshold was set at
166 50pA to capture all significant peaks while avoiding inclusion of small peaks likely representing
167 contamination from the hexane solvent. Note that this resulted in the merger of the closely eluting
168 caryophyllene oxide and ethyl laurate peaks, which were removed from the analyses of compound type
169 and compound modification impact on GC-EAD responses as a result.

Data loading and initial processing

EAD and FID waveforms were loaded as CSV files into R (R Core Team, 2024; see Appendix for example script). GC-FID onsets and offsets determined in the previous section were also loaded. As the EAD setup samples at a rate of 6000 datapoints per second and does not export a timestamp to its CSV file, a new time variable was defined corresponding to the sampling rate. The GC-FID injection time and a 20 minute endpoint (after all tested peaks had eluted) relative to the EAD time were also defined, as were EAD timestamps corresponding to each GC-FID peak onset and offset. As injections were done manually for the GC-FID, there was occasionally a slight disparity between GC-FID peak onsets and the onset of a significant-seeming EAD pattern, and so a user-defined offset was determined by comparing the most significant EAD pattern onset with the previously determined corresponding GC-FID peak onset.

First Forward Difference (FFD) processing and significant FFD value identification

Using the diff() function in base R (specifically as 'diff(EADamplitude,lag=1,differences=1)'), the First Forward Difference of the antennal amplitude was calculated from the input antennal amplitude data. The FFD vector was subsetted into five windows of four minutes each for a total GC-FID run of 20 minutes. For each window, the average and standard deviation of the FFD within that window was calculated. A significance threshold of 3σ above or below the average was calculated, and data points that exceeded the significance threshold for their respective windows were marked as significant.

Calculating the number of significant FFD hits per second for each identified GC-FID peak

For each GC-FID peak identified, the number of significant hits as defined above within its onset-offset window was determined using the 'between' function in the R package dplyr (Wickham et al., 2023). This value was divided by the length of the GC-FID peak in seconds (offset time minus onset time) to get the significant hit frequency of hits/second. As EAD data can have significant background noise (Schiestl

192 and Marion-Poll, 2002), we defined the frequency of the initial hexane peak 1 as the threshold for
193 whether a peak attracted significant EAD activity: if the frequency for a given GC-FID peak exceeded the
194 frequency of the hexane peak, it was counted as significant, and otherwise considered insignificant.

195 *Statistical analysis*

196 A "significant response" across each species-sex/caste category (female *M. sexta*, male *M. sexta*, and
197 worker *B. terrestris*) was identified for compounds where >50% of individuals tested showed a
198 significant hit frequency greater than that of the initial large solvent (hexane) peak.

199 To investigate whether the results were repeatable between moths, we calculated the Pearson
200 correlation coefficient between the significant hit frequency for each pair of moths within a sex using
201 cor.test() in R. We repeated this test with a binary vector (1 = GC-FID peak had a significant hit
202 frequency; 0 = it did not). We then used lm() in R to calculate the slope and intercept of the relationship
203 between each pair of moths within a sex. Finally, we used lm() and emmeans() (Lenth, 2024) to
204 determine whether compound skeletal or modification type had an influence on significant hit
205 frequency within each species-sex/caste group (female *M. sexta*, male *M. sexta*, and worker *B.*
206 *terrestris*).

207 **Results**

208 *The new method allows quantitative analysis and avoids baseline drift*

209 Traditional GC-EAD analysis is done in a visual qualitative fashion by comparing GC-FID peak timings with
210 EAD readouts by eye (e.g. comparison of **Figure 1A** with **Figure 1B**, see also **Figures S1-S15**). It also
211 suffers from baseline drift and signal noise, two phenomena that make it difficult to detect whether an
212 EAD data point is indeed a "significant" response to a GC-provided stimulus or not (Schiestl and Marion-
213 Poll, 2002). As the antenna degrades over time and accommodates in the long term to compound

214 stimuli (see Byers et al., 2020), antennal responsiveness and the ability to detect significant responses by
215 eye also degrades and the noise level often increases, also impeding qualitative visual comparison of GC-
216 FID and EAD data.

217 Here we present a method that avoids the otherwise significant biases inherent in GC-EAD *qualitative*
218 analysis methodology by the use of a signal processing technique, the *first forward difference* (FFD), to
219 remove the influence of baseline drift, essentially flattening out the EAD signal and replacing its
220 waveform with a series of individual spikes (**Figure 1C**, see also **Figures S16-S30**). A significance
221 threshold may then be set for each window (horizontal lines in **Figure 1C** and **Figures S16-S30**, both
222 showing ± 3 standard deviations above/below the mean for that window), and the frequency of spikes
223 with an amplitude over that significance threshold during a given GC-FID peak may be easily calculated
224 and analysed using traditional statistical methods. This results in a *quantitative* rather than *qualitative*
225 assessment of whether a GC-EAD peak is physiologically active or not.

226 *GC-FID peaks with significant EAD responses*

227 **Table 1** shows the list of compounds tested in the EAD test mixture along with their response rates for
228 female and male moths and worker bumble bees and how many individuals of each had a response rate
229 over that of the main hexane peak. In general, female moths were most responsive to fatty acid-derived
230 compounds (FADs, red in **Figure 2A**, see also **Figure S31**) and to some but not all aromatic volatiles (blue
231 in **Figure 2A**). Responses to terpenoids were more mixed (green in **Figure 2A**), with some compounds
232 (e.g. (1S)-(-)-verbenone) eliciting significant responses, while others (e.g. camphene) eliciting no
233 significant responses at all. The sole tested nitrogenous compound, indole (yellow in **Figure 2A**), did not
234 elicit significant responses in female moths. Overall, female moths responded significantly to 15 of the
235 31 tested compounds (excluding the hexane peaks). Male moths were far less responsive than females
236 to the majority of compounds, with overall significant responses to only four of the 31 tested

compounds (excluding the hexane peaks) and with no significant difference in response between compound skeleton types (**Figure S31A,B**). Worker bumble bees responded significantly to 10 of the 31 compounds. Bees responded equally to all types of compound classes (**Figure S31C**), although there was no response to indole, the sole nitrogenous compound tested, in any bee.

Compounds containing oxygenated modifications seemed to produce the highest level of response in female moths, but not in males (**Figure S31D,E**). Females responded significantly to 6 of 10 presented esters (analyzed without the co-eluting peaks ethyl laurate and caryophyllene oxide), 2/6 alcohols, 3/3 aldehydes, and 3/3 ketones, but not to any of the five presented unmodified monoterpenes ((1R)-α-pinene, camphene, and R-(+)-limonene) or sesquiterpenes (caryophyllene, α-humulene), both groups of which lack oxygen atoms. Male moths responded to 1 of 10 esters, 0/6 alcohols, 1/3 aldehydes, and 0/3 ketones but, similarly to females, only responded to 1/5 of the unmodified (mono/sesqui)terpenes. Both females and males also responded significantly overall to the "Unknown 2" compound, whose identity remains unclear despite GC-MS analysis but which likely also is oxygenated based on tentative NIST library matches. For bees, there was no difference in response based on compound modification, oxygenated or otherwise (**Figure S31F**). Bees responded to 3 of 10 esters, 3/6 alcohols, 1/3 aldehydes, 2/3 ketones, and 1/5 unmodified (mono/sesqui)terpenes.

Repeatability within moth sexes and comparisons between sexes

Responses to non-hexane compounds were highly repeatable between female moths (**Figure S32A**, Pearson's correlation coefficient range 0.265 to 0.637, all $P < 3x10^{-7}$), but not between male moths (**Figure S32B**, Pearson's correlation coefficient range -0.238 to -0.054, all $P \geq 0.20$). In general, male moths were substantially less responsive to non-hexane chemical stimuli in this experiment, as represented by their much lower average significant hit frequency (**Table 1,** range 0-4.000 significant hits/second) versus females (range 0-21.919 significant hits/second). Responses between males and

260 females were generally uncorrelated (Pearson's r = 0.073, *P* = 0.696, **Figure S33**). The only two
261 compounds that resulted in significant hit frequencies in ≥ 2/3 of moths of both sexes were
262 phenylacetaldehyde and "Unknown 1".

263 *Repeatability and feasibility across species and comparison to existing literature*

264 To ensure that our results were not simply due to some property of *Manduca sexta* antennae, we also
265 analysed the same stimulus mixture with four workers of the polylectic (wide dietary breadth, Rasmont
266 et al., 2008) bumble bee *Bombus terrestris audax* using the same data collection and analysis
267 methodologies. We found no compounds that were significant in >50% of tested individuals across all
268 three species-sex/caste combinations (female hawkmoths, male hawkmoths, worker bumble bees).
269 However, we did identify a short list of compounds that were significant in >50% of tested individuals for
270 both female *Manduca sexta* and worker *Bombus terrestris audax*, namely benzaldehyde, (*Z*)-3-hexenyl
271 acetate, acetophenone, (1S)-(-)-verbenone, and geraniol (**Table 1**). Results from individuals of all three
272 groups are presented in **Figure 3**. We also observed overall weaker responses in bumble bees compared
273 with hawkmoths, despite robust responses to single puffs of the EAD stimulus mixture. The signal to
274 noise threshold in the bumble bee antennae appears to be weaker than in the hawkmoths (as shown by
275 much lower first forward difference significance thresholds and a lower range of significant average hit
276 frequencies (range 0-1.752 hits/second compared with hawkmoth values of 0-21.919 and 0-4.000 for
277 females and males respectively). We separately conducted a somewhat haphazard literature review to
278 identify whether our tested compounds had previously been identified as electroantennographically
279 active in either *Manduca sexta* or *Bombus terrestris*, with the results shown in **Table S3**. In general,
280 many of our tested compounds had not previously been tested with GC-EAD or puff-based
281 electroantennography, and those that had been often produced results that differed between studies,
282 further suggesting the need for a quantitative analysis technique such as the one presented here.

CONCLUSIONS

In this Protocol Note, we have presented a novel signal processing-based methodology for use with gas chromatography-coupled electroantennographic detection (GC-EAD) measurements of insect antennae in response to plant volatiles. Female *Manduca sexta* hawkmoths were responsive to approximately half of tested compounds in a highly repeatable fashion, while male hawkmoths of the same species responded to only about one eighth of tested compounds and were far less repeatable between individuals. In female moths, compound skeleton type (FAD, aromatic, terpenoid, or nitrogenous) and modification (none, alcohol, aldehyde, ester, ketone, or oxide) played a significant role in response patterns, while in male moths this was only true for modification type but not compound skeleton type. Worker *Bombus terrestris audax* bumble bees showed an intermediate level of response, with significant responses to one third of tested compounds, and with no difference in response rate relative to compound type or modification.

Our haphazard literature review to identify compounds that had previously been identified as electroantennographically active in either *Manduca sexta* or *Bombus terrestris* identified large gaps in our knowledge, as well as discrepancies between studies in whether compounds were designated as physiologically active or not in the antenna**.** Consistent with our findings, some compounds were significant in some studies but not others, emphasizing the importance of testing stimuli on multiple individuals of a given species rather than relying on one or two individuals, given that we found variability even between individuals of the same commercial bumble bee hive and hawkmoth colony. The weaker responses seen in bumble bees further necessitate the use of this new technique to ensure unbiased assessment of electroantennographic detection studies in potentially less strongly responsive insects.

By studying responses to common plant volatiles in two model pollinator species, we have shown that responses are species- and sex-specific, and that it will be difficult to generalize results from one pollinator group to another. Further study will be needed to determine whether results can be generalized within a genus of pollinators, or between those that are generalists and specialists. However, at least one species of pollinating insect was responsive to the majority of the plant compounds tested here. Interestingly, there was no significant GC-EAD response in any species tested to the monoterpene (R)-(+)-limonene. Limonene is the most widely distributed floral volatile in terms of angiosperm families in which it is found (Schiestl, 2010), and is known to provoke a responsive in the antennal lobe of *Bombus vosnesenskii* (Byers et al., 2014a). We also did not see any response to the common nitrogenous volatile indole. The latter is utilized in many dung-mimicking brood site-deceptive pollination systems (Jürgens et al., 2013), which does not include the insects tested here.

The methods introduced make use of existing typical GC-EAD recording setups and methods, differing only in the downstream processing of collected GC-EAD data, and thus may be used with both existing as well as novel datasets. In addition, they are simple to execute, requiring only minimal implementation in R or another statistical processing language (example code is provided in the Appendix). This study provides the first truly quantitative analysis of insect responses to common plant volatiles using GC-EAD, and paves the way forward for future GC-EAD analyses to incorporate quantitative, unbiased data processing methodologies to determine whether plant volatiles are physiologically active to interacting insect species.

AUTHOR CONTRIBUTIONS

KJRPB conceived the project, designed and executed the data collection and data analysis methodology, and wrote the manuscript. RNJ contributed significant signal processing methodology ideas and

329 expertise to the data analysis methodology development. Both authors approved the final version of the
330 manuscript.

331

332 **ACKNOWLEDGMENTS**

333 The authors thank the John Innes Centre (JIC) Entomology platform for maintaining the *Manduca sexta*
334 and *Bombus terrestris* colonies and the JIC Metabolomics platform for GC-MS maintenance. Funding was
335 provided by the UK Biotechnology and Biological Sciences Research Council Institute Strategic
336 Programmes (Harnessing Biosynthesis for Sustainable Food and Health (HBio) (grant no. BB/X01097X/1)
337 and Building Robustness in Crops (BRiC) (grant no. BB/X01102X/1)). We also thank anonymous
338 reviewers and the editor for their helpful suggestions.

339

340 **DATA AVAILABILITY**

341 All raw data (GC-FID CSV files, electroantennography CSV files, and relevant GC-FID injection
342 timestamps) are available on FigShare (doi: XXX). Scripts referenced in the main text and appendices are
343 also available on GitHub ().

344

345 **REFERENCES**

346 Akers, R. P., and W. M. Getz. 1993. Response of olfactory receptor neurons in honeybees to odorants
347 and their binary mixtures. *Journal of Comparative Physiology A* 173: 169-185.

348 Byers, K. J. R. P. 2021. "As if they discovered it by the scent": improving our understanding of the
349 chemical ecology, evolution, and genetics of floral scent and its role in pollination. *American Journal of*
350 *Botany* 108: 729-731.

Byers, K. J. R. P. 2023. Reducing eggs on eggplant: a common naturally emitted plant volatile could replace insecticides in the 'king of vegetables'. *New Phytologist* 240: 915-917.

Byers, K. J. R. P., H. D. Bradshaw, Jr., and J. A. Riffell. 2014a. Three floral volatiles contribute to differential pollinator attraction in monkeyflowers (*Mimulus*). *Journal of Experimental Biology* 217: 614-623.

Byers, K. J. R. P., J. P. Vela, F. Peng, J. A. Riffell, and H. D. Bradshaw, Jr. 2014b. Floral volatile alleles can contribute to pollinator-mediated reproductive isolation in monkeyflowers (*Mimulus*). *The Plant Journal* 80: 1031-1042.

Byers, K. J. R. P., K. Darragh, J. Musgrove, D. Abondano Almeida, S. Fernanda Garza, I. A. Warren, P. M. Rastas, M. Kučka, Y. F. Chan, R. M. Merrill, S. Schulz, W. O. McMillan, and C. D. Jiggins. 2020. A major locus controls a biologically active pheromone component in *Heliconius melpomene*. *Evolution* 74: 349-364.

Clavijo McCormick, A., J. Gershenzon, and S. B. Unsicker. 2014. Little peaks with big effects: establishing the role of minor plant volatiles in plant-insect interactions. *Plant, Cell & Environment* 37: 1836-1844.

Dixon, R. A., and A. J. Dickinson. 2024. A century of studying plant secondary metabolism - From "what" to "where, how, and why?" *Plant Physiology* 195: 48-66.

Gallage, N. J., E. H. Hansen, R. Kannangara, C. E. Olsen, M. S. Motawia, K. Jørgensen, I. Holme, K. Hebelstrup, M. Grisoni, and B. L. Møller. 2014. Vanillin formation from ferulic acid in *Vanilla planifolia* is catalysed by a single enzyme. *Nature Communications* 5: 4037.

Ghosh, R., D. Metze, S. Sant, M. Shaikh, A. Deshpande, D. M. Firake, and S. Pandit. 2023. Chemical ecology of Himalayan eggplant variety's antixenosis: identification of geraniol as an oviposition deterrent against the eggplant shoot and fruit borer. *New Phytologist* 240: 1259-1274.

Jürgens, A., S.-L. Wee, A. Shuttleworth, and S. D. Johnson. 2013. Chemical mimicry of insect oviposition sites: a global analysis of convergence in angiosperms. *Ecology Letters* 16: 1157-1167.

Kessler, A., and A. Kalske. 2018. Plant secondary metabolite diversity and species interactions. *Annual Review of Ecology, Evolution, and Systematics* 49: 115-138.

Lenth, R. 2024. Version 1.10.0 (2024-04-06). emmeans: Estimated Marginal Means, aka Least-Squares Means. Website https://CRAN.R-project.org/package=emmeans [accessed 1 December 2024].

Peakall, R., D. Ebert, J. Poldy, R. A. Barrow, W. Francke, C. C. Bower, and F. P. Schiestl. 2010. Pollinator specificity, floral odour chemistry and the phylogeny of Australian sexually deceptive *Chiloglottis* orchids: implications for pollinator-driven speciation. *New Phytologist* 188: 437-450.

Perl. 2023. Version 5.38.2 (2023-11-29). Perl. Website https://www.perl.org [accessed 1 December 2024].

Pickett, J. A., and Z. R. Khan. 2016. Plant volatile-mediated signalling and its application in agriculture: successes and challenges. *New Phytologist* 212: 856-870.

R Foundation for Statistical Computing. 2024. Version 4.3.2 (2023-10-31). The R Foundation for Statistical Computing, Vienna Austria. Website http://www.r-project.org/ [accessed 1 December 2024].

Raguso, R. A. 2008. Wake up and smell the roses: the ecology and evolution of floral scent. *Annual Review of Ecology, Evolution, and Systematics* 39: 549-569.

Rasmont P., A. Coppee, D. Michez, and T. De Meulemeester. 2008. An overview of the *Bombus terrestris* (L. 1758) subspecies (Hymenoptera: Apidae). *Annales de la Société entomologique de France (N.S.)* 44: 243-250.

Schiestl, F. P. 2010. The evolution of floral scent and insect chemical communication. *Ecology Letters* 13: 643-656.

Schiestl, F. P., and F. Marion-Poll. 2002. Detection of physiologically active flower volatiles using gas chromatography coupled with electroantennography. In J. F. Jackson, H. F. Linskens [eds.], Analysis of Taste and Aroma, Molecular Methods of Plant Analysis, volume 21, 173-198. Springer, Berlin, Heidelberg, Germany.

Schneider, D. 1957. Electrophysiological investigation on the antennal receptors of the silk moth during chemical and mechanical stimulation. *Experientia* 13: 89-91.

Wickham, H., R. François, L. Henry, K. Müller, and D. Vaughan. 2023. Version 1.1.4 (2023-11-17). dplyr: A Grammar of Data Manipulation. Website https://CRAN.R-project.org/package=dplyr [accessed 1 December 2024].

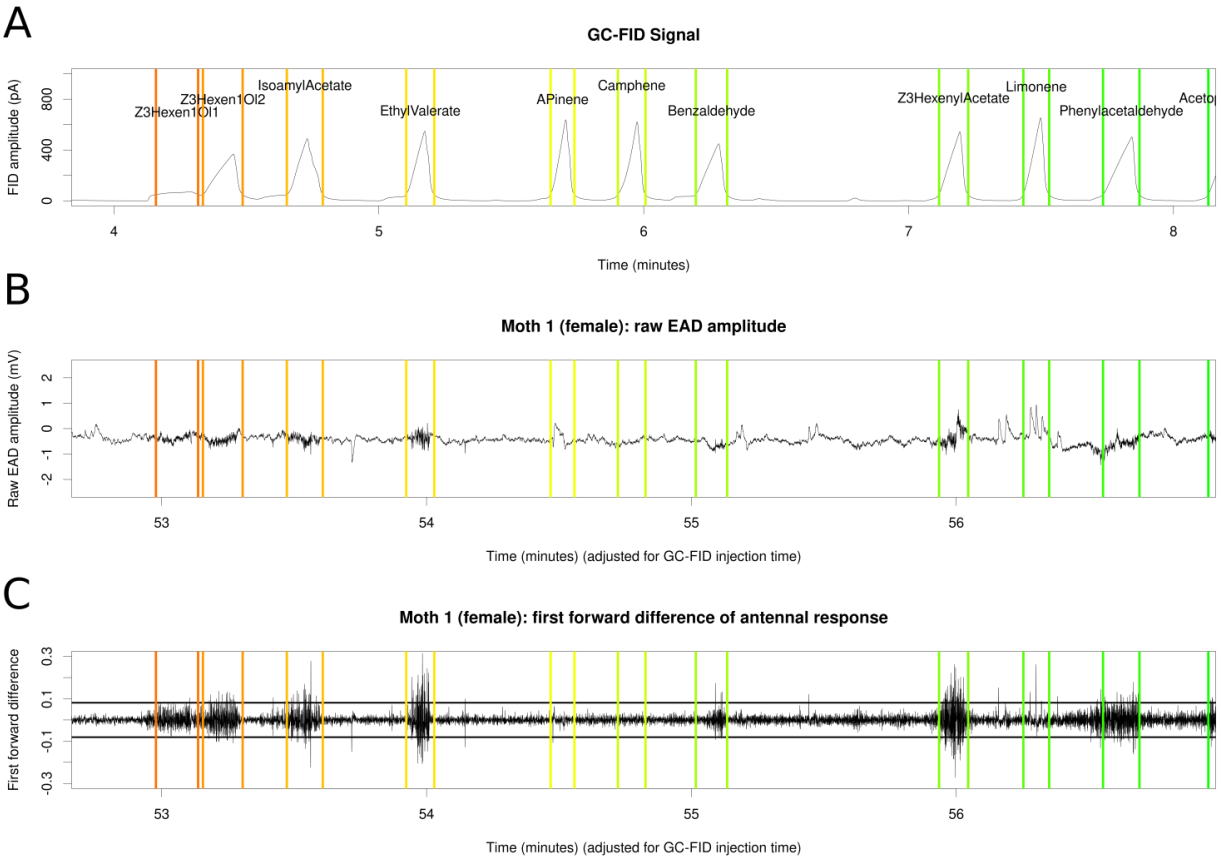
APPENDIX

The methods used in this Protocol Note may be found in detail in the Methods section, but included below are two computational code examples that users may find helpful:

1. The Perl script used to identify peak onsets and offsets in GC-FID CSV data;

2. Example R code to process GC-EAD data and identify GC-FID peaks with significant hit rates over the hexane peak background level.

411 **FIGURES**



412

413 **Figure 1:** Comparison of raw EAD and first forward difference-processed EAD traces from a single female
414 *Manduca sexta* hawkmoth antenna between 4-8 minutes of GC elapsed time. (A) GC-FID trace annotated
415 with automatically detected peak starts and ends (colored vertical lines) and peak identities. (B) Raw
416 EAD amplitude annotated with GC-FID peak starts and ends. (C) First forward difference-processed EAD
417 amplitude annotated with GC-FID peak starts and ends. Note the "wandering baseline" in (B) that is
418 absent in (C).

419

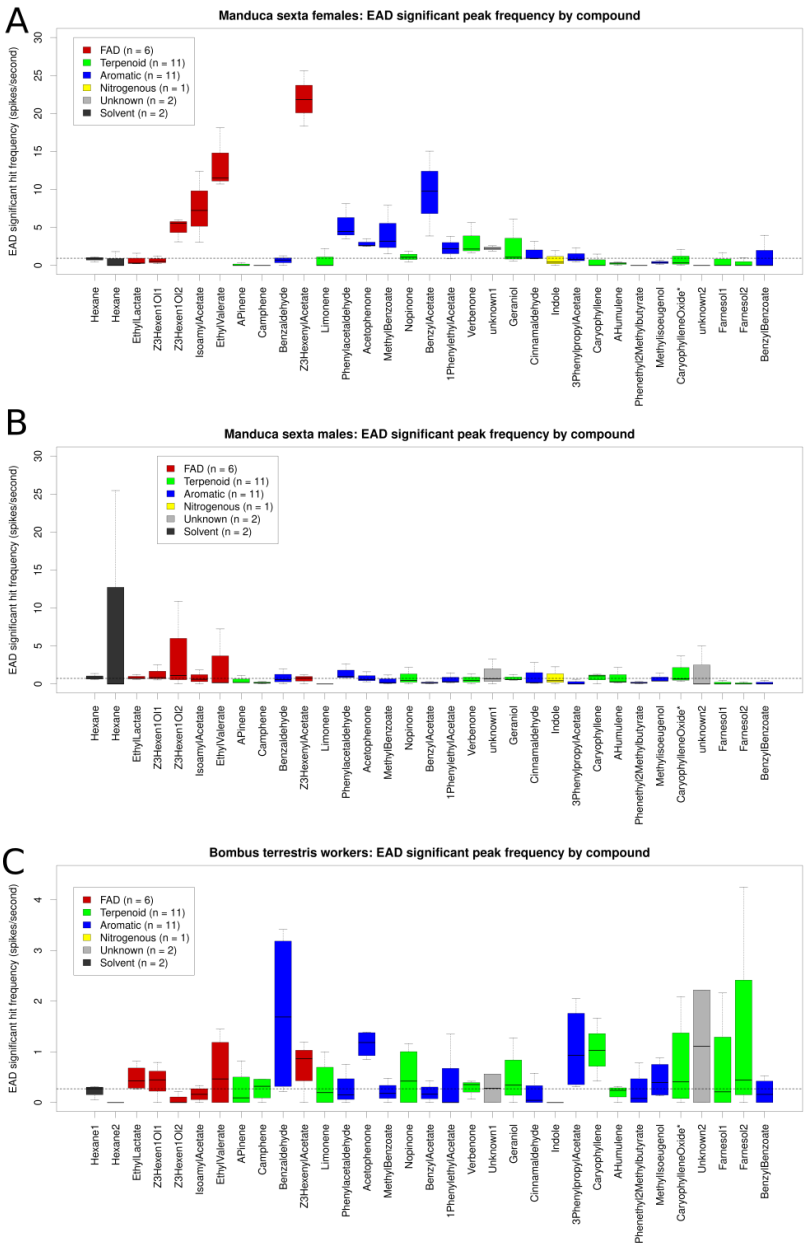


Figure 2: First forward difference-processed EAD responses to 33 compounds from the test mixture, colored by compound type (note that the Y-axis scaling differs for (C)). Note for caryophyllene oxide*: this box includes responses to ethyl laurate, as the peaks were merged in the GC-FID annotation thresholding step so are treated as a single peak by the analysis. (A) Data from three *Manduca sexta* females. (B) Data from three *Manduca sexta* males. (C) Data from four *Bombus terrestris audax* workers.

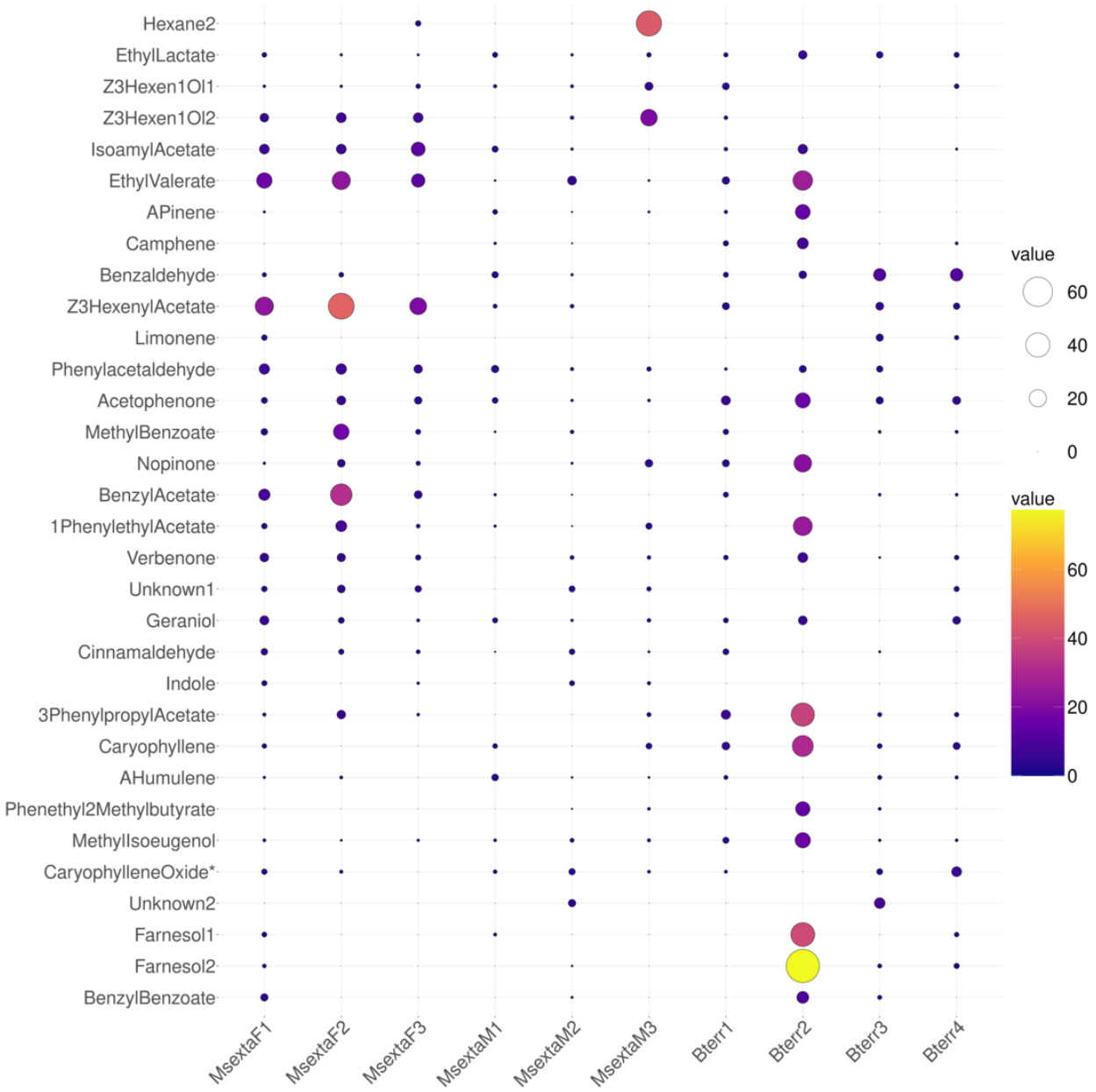


Figure 3: Antennal responses of three female *Manduca sexta* ("MsextaF[1-3]"), three male *Manduca sexta* ("MsextaM[1-3]"), and four *Bombus terrestris audax* workers ("Bterr[1-4]") expressed as a ratio of the significant hit frequency of each compound to that of the significant hit frequency of the initial hexane solvent peak ("value"), with warmer colors and larger circles denoting higher ratios and blue and smaller circles denoting lower ratios. Absent points indicate the specific peak was not detected in the GC-FID run for a given insect.

436 **TABLES**

437 **Table 1:** Compounds used in the EAD test mixture with details of their retention time, compound
438 skeleton type and subtype or modification categorization, and average significant hit frequency (spikes
439 per second, bold text indicates higher than first hexane peak frequency for that species and sex/caste)
440 for female and male *Manduca sexta* hawkmoths and worker *Bombus terrestris audax* bumble bees. The
441 ratio after the significant hit frequency indicates the number of samples where the given compound
442 elicited a greater response than the main hexane peak versus the total number of samples with that
443 peak present (bold: >50% of samples reacted to that peak for a given species-sex/caste combination).
444 Unknown superscripts: 1[A]: an unclear oxygenated cyclic molecule from GC-MS results; 2[B]: likely an
445 unknown sesquiterpenoid oxide from GC-MS analysis results.

Retention time (GC-FID, minutes)	**Compound identity**	**Skeleton**	**Modifications / subcategory**	**Average *Manduca sexta* female response (significant hit frequency, spikes/second)**	**Average *Manduca sexta* male response (significant hit frequency, spikes/second)**	**Average *Bombus terrestris audax* worker response (significant hit frequency, spikes/second)**
1.973	Hexane, main peak	FAD	None	0.843 (NA)	0.896 (NA)	0.229 (NA)
2.687	Hexane, small peak	FAD	None	0.606 (1/3)	8.485 (1/3)	0 (0/1)
3.755	Ethyl L-lactate	FAD	Ester	0.700 (1/3)	**0.867 (2/3)**	**0.483 (4/4)**
4.291	(*Z*)-3-hexen-1-ol, shoulder peak	FAD	Alcohol	0.663 (1/3)	1.292 (1/3)	**0.413 (2/3)**
4.458	(*Z*)-3-hexen-1-ol, main peak	FAD	Alcohol	**4.889 (3/3)**	4.000 (1/3)	0.055 (0/4)
4.741	Isoamyl acetate	FAD	Ester	**7.566 (3/3)**	0.818 (1/3)	0.165 (1/4)
5.184	Ethyl valerate	FAD	Ester	**13.438 (3/3)**	2.519 (1/3)	0.594 (2/4)
5.716	(1R)-(+)-α-pinene	Terpenoid	Monoterpene	0.123 (0/3)	0.494 (1/3)	0.251 (1/4)
5.985	Camphene	Terpenoid	Monoterpene	0 (0/3)	0.160 (0/3)	0.276 (2/4)
6.292	Benzaldehyde	Aromatic	Aldehyde	**0.657 (2/3)**	0.845 (1/3)	**1.752 (4/4)**
7.200	(*Z*)-3-hexenyl acetate	FAD	Ester	**21.919 (3/3)**	0.657 (1/3)	**0.730 (3/4)**
7.505	(R)-(+)-limonene	Terpenoid	Monoterpene	0.741 (1/3)	0 (0/3)	0.348 (2/4)
7.847	Phenylacetaldehyde	Aromatic	Aldehyde	**5.374 (3/3)**	**1.455 (2/3)**	0.265 (2/4)
8.222	Acetophenone	Aromatic	Ketone	**2.871 (3/3)**	0.827 (1/3)	**1.152 (4/4)**
8.691	Methyl benzoate	Aromatic	Ester	**4.202 (3/3)**	0.448 (0/3)	0.210 (1/4)
9.355	(1R)-(+)-nopinone	Terpenoid	Monoterpenoid ketone	**1.137 (2/3)**	0.879 (1/3)	0.502 (2/4)
9.818	Benzyl acetate	Aromatic	Ester	**9.561 (3/3)**	0.155 (0/3)	0.190 (1/4)
10.279	1-phenylethyl	Aromatic	Ester	**2.304 (2/3)**	0.637 (1/3)	0.337 (1/4)

	acetate					
10.511	(1S)-(-)-verbenone	Terpenoid	Monoterpenoid ketone	**3.154 (3/3)**	0.599 (0/3)	**0.299 (3/4)**
10.609	Unknown 1[A]	Unclear	Unclear	**2.241 (3/3)**	**1.322 (2/3)**	0.282 (1/2)
11.240	Geraniol	Terpenoid	Monoterpenoid alcohol	**2.583 (2/3)**	0.778 (1/3)	**0.491 (3/4)**
11.459	(*E*)-cinnamaldehyde	Aromatic	Aldehyde	**1.645 (2/3)**	1.021 (1/3)	0.167 (1/4)
11.777	Indole	Nitrogenous	Aromatic	0.788 (1/3)	0.896 (1/3)	0 (0/4)
12.691	3-phenylpropyl acetate	Aromatic	Ester	1.189 (1/3)	0.207 (1/3)	**1.057 (4/4)**
13.254	β-caryophyllene	Terpenoid	Sesquiterpene	0.496 (1/3)	**0.780 (2/3)**	**1.038 (4/4)**
13.683	α -humulene	Terpenoid	Sesquiterpene	0.240 (0/3)	0.863 (1/3)	0.200 (2/4)
14.111	Phenethyl 2-methylbutyrate	Aromatic	Ester	0 (0/3)	0.159 (0/3)	0.235 (1/4)
14.285	Methyl isoeugenol	Aromatic	Alcohol	0.387 (0/3)	0.731 (1/3)	0.450 (2/4)
15.265	Caryophyllene oxide + ethyl laurate (peaks merged by FID peak identification script)	Terpenoid + FAD	Sesquiterpenoid oxide + ester	0.816 (1/3)	1.551 (1/3)	0.727 (2/4)
15.476	Unknown 2[B]	Terpenoid	Unknown sesquiterpenoid oxide	0 (0/3)	1.667 (1/3)	1.111 (1/2)
16.474	Farnesol, isomer 1	Terpenoid	Sesquiterpenoid alcohol	0.564 (1/3)	0.154 (0/3)	0.646 (2/4)
16.734	Farnesol, isomer 2	Terpenoid	Sesquiterpenoid alcohol	0.333 (0/3)	0.095 (0/3)	**1.284 (3/4)**
17.189	Benzyl benzoate	Aromatic	Ester	1.323 (1/3)	0.153 (0/3)	0.212 (2/4)

446

447

SUPPORTING INFORMATION: SUPPLEMENTAL FIGURE LEGENDS AND TABLE

Figures S1-S5: Raw EAD amplitude for three female hawkmoths. Top pane: GC-FID trace annotated with peak starts and ends (colored vertical lines) and peak identities. Bottom three panes: individual female hawkmoth EAD traces annotated with peak starts and ends (colored vertical lines). Figure S1: 0-4 minutes of GC-FID elapsed time. Figure S2: 4-8 minutes. Figure S3: 8-12 minutes. Figure S4: 12-16 minutes. Figure S5: 16-20 minutes. In Figure S5, the large peaks in the second two EAD traces after the benzyl benzoate interval are mixture puffs to test antennal viability.

Figures S6-S10: Raw EAD amplitude for three male hawkmoths. Top pane: GC-FID trace annotated with peak starts and ends (colored vertical lines) and peak identities. Bottom three panes: individual male hawkmoth EAD traces annotated with peak starts and ends (colored vertical lines). Figure S1: 0-4 minutes of GC-FID elapsed time. Figure S2: 4-8 minutes. Figure S3: 8-12 minutes. Figure S4: 12-16 minutes. Figure S5: 16-20 minutes. In Figure S10, the large peaks in the EAD traces after the benzyl benzoate interval are mixture puffs to test antennal viability.

Figures S11-S15: Raw EAD amplitude for four bumble bee workers. Top pane: GC-FID trace annotated with peak starts and ends (colored vertical lines) and peak identities. Bottom four panes: individual bumble bee worker EAD traces annotated with peak starts and ends (colored vertical lines). Figure S11: 0-4 minutes of GC-FID elapsed time. Figure S12: 4-8 minutes. Figure S13: 8-12 minutes. Figure S14: 12-16 minutes. Figure S15: 16-20 minutes.

Figures S16-S20: First forward difference-processed EAD amplitude for three female hawkmoths. Top pane: GC-FID trace annotated with peak starts and ends (colored vertical lines) and peak identities. Bottom three panes: individual female hawkmoth first forward difference-processed EAD traces annotated with peak starts and ends (colored vertical lines) and significant peak hit frequency (spikes/second), with red indicating frequency over the median hexane solvent threshold. Figure S16: 0-4 minutes of GC-FID elapsed time. Figure S17: 4-8 minutes. Figure S18: 8-12 minutes. Figure S19: 12-16 minutes. Figure S20: 16-20 minutes. In Figure S20, the large peaks in the second two EAD traces after the benzyl benzoate interval are mixture puffs to test antennal viability.

Figures S21-S25: First forward difference-processed EAD amplitude for three male hawkmoths. Top pane: GC-FID trace annotated with peak starts and ends (colored vertical lines) and peak identities. Bottom three panes: individual male hawkmoth first forward difference-processed EAD traces annotated with peak starts and ends (colored vertical lines) and significant peak hit frequency (spikes/second), with red indicating frequency over the median hexane solvent threshold. Figure S21: 0-4 minutes of GC-FID elapsed time. Figure S22: 4-8 minutes. Figure S23: 8-12 minutes. Figure S24: 12-16 minutes. Figure S25: 16-20 minutes. In Figure S25, the large peaks in the EAD traces after the benzyl benzoate interval are mixture puffs to test antennal viability.

Figures S26-S30: First forward difference-processed EAD amplitude for four bumble bee workers. Top pane: GC-FID trace annotated with peak starts and ends (colored vertical lines) and peak identities. Bottom three panes: individual bumble bee worker first forward difference-processed EAD traces annotated with peak starts and ends (colored vertical lines) and significant peak hit frequency

(spikes/second), with red indicating frequency over the median hexane solvent threshold. Figure S16: 0-4 minutes of GC-FID elapsed time. Figure S17: 4-8 minutes. Figure S18: 8-12 minutes. Figure S19: 12-16 minutes. Figure S20: 16-20 minutes.

Figure S31: First forward difference-processed EAD significant peak hit frequency (spikes/second) across (A,B,C) four chemical skeleton classes (FADs, terpenoids, aromatics, and nitrogenous compounds) and (D,E,F) six compound modification types tested (none, alcohols, aldehydes, esters, and ketones) for (A,D) three *Manduca sexta* females, (B,E) three *Manduca sexta* males, and (C,F) four *Bombus terrestris audax* workers.

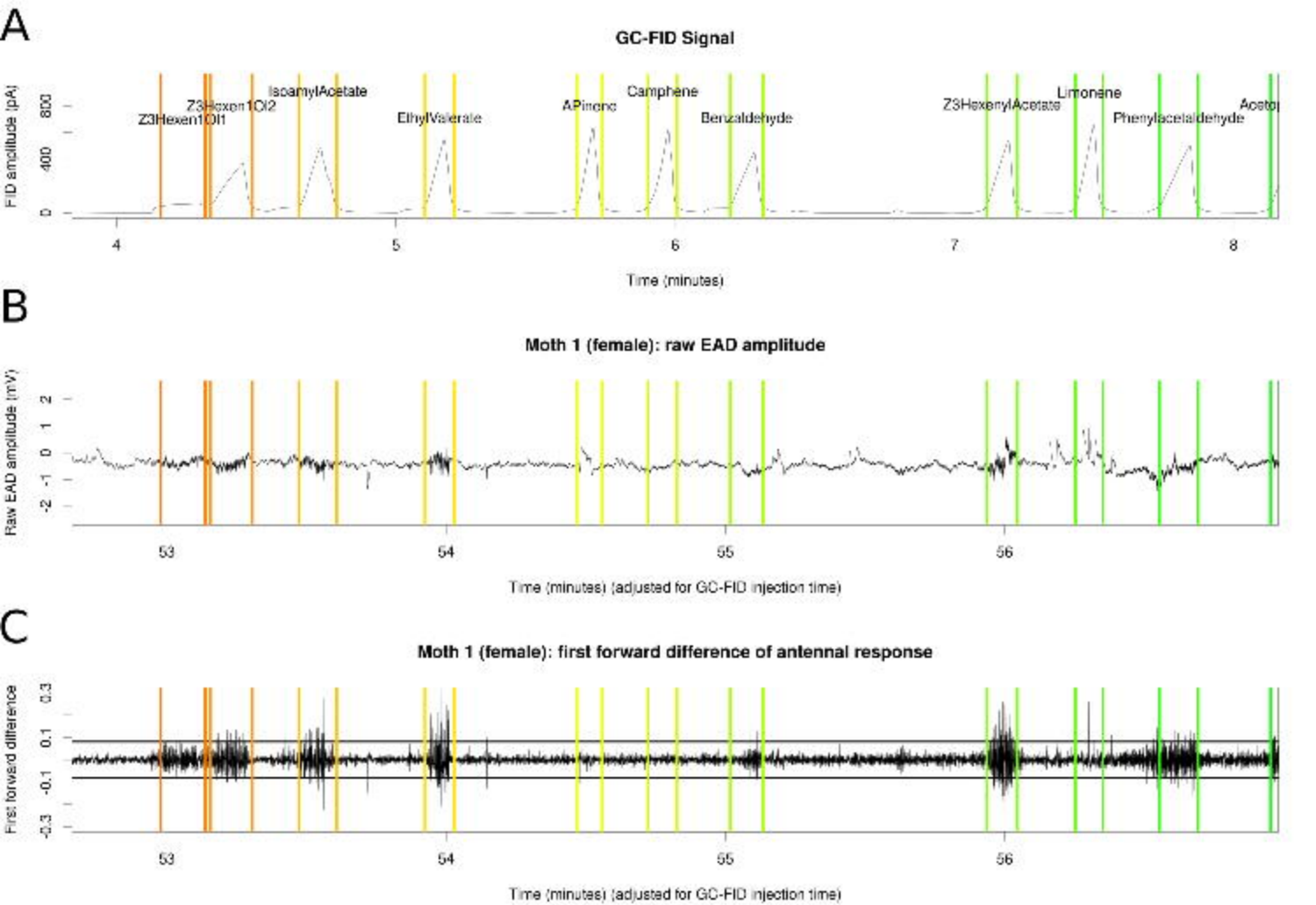
Figure S32: Inter-sex EAD response (repeatability) comparisons between the average values of male and female *Manduca sexta* hawkmoths for each GC-FID peak.

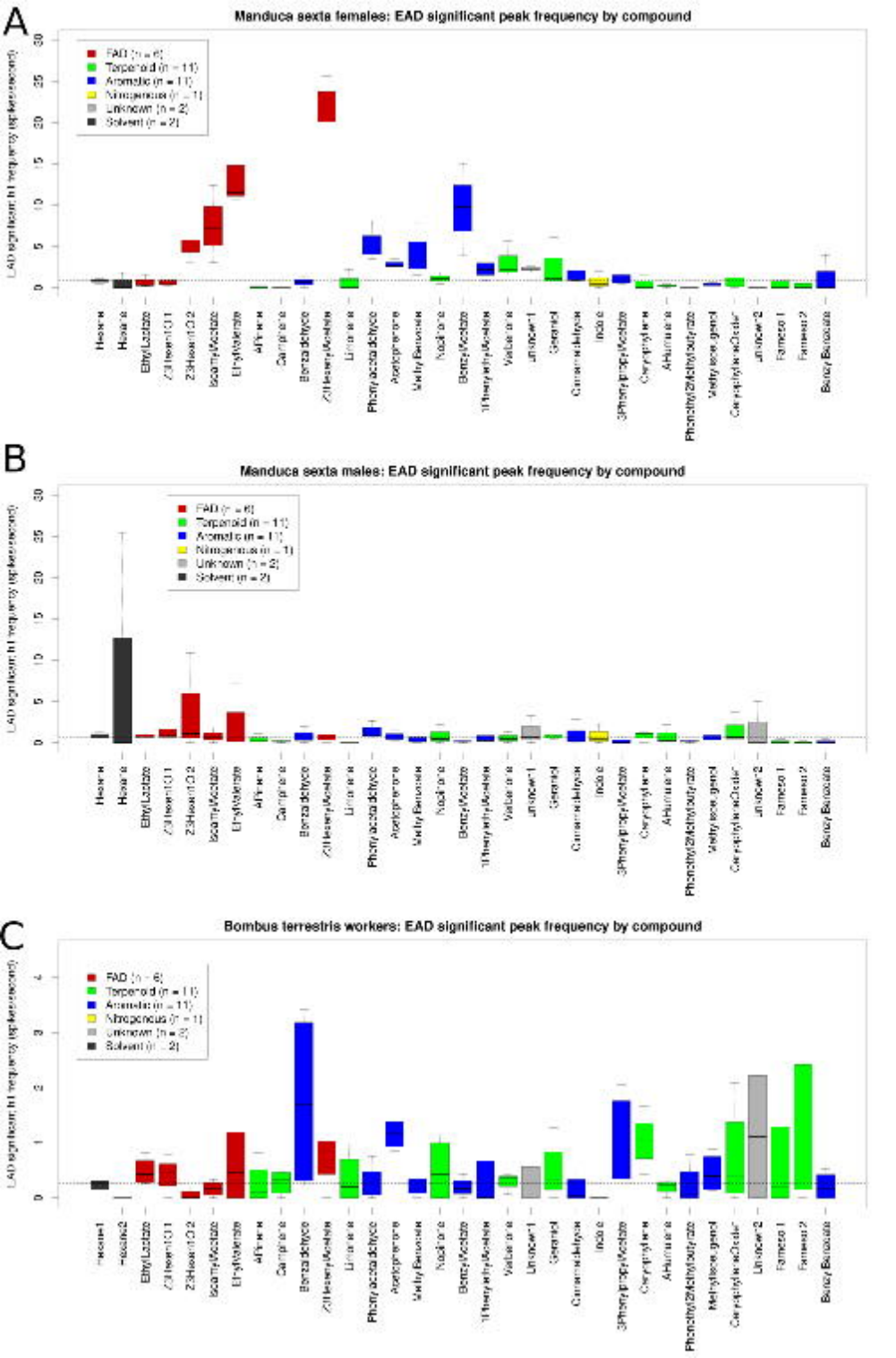
Figure S33: Inter-individual EAD response (repeatability) comparisons between *Manduca sexta* hawkmoths of the same sex. (A) All pairwise comparisons between three females. (B) All pairwise comparisons between three males.

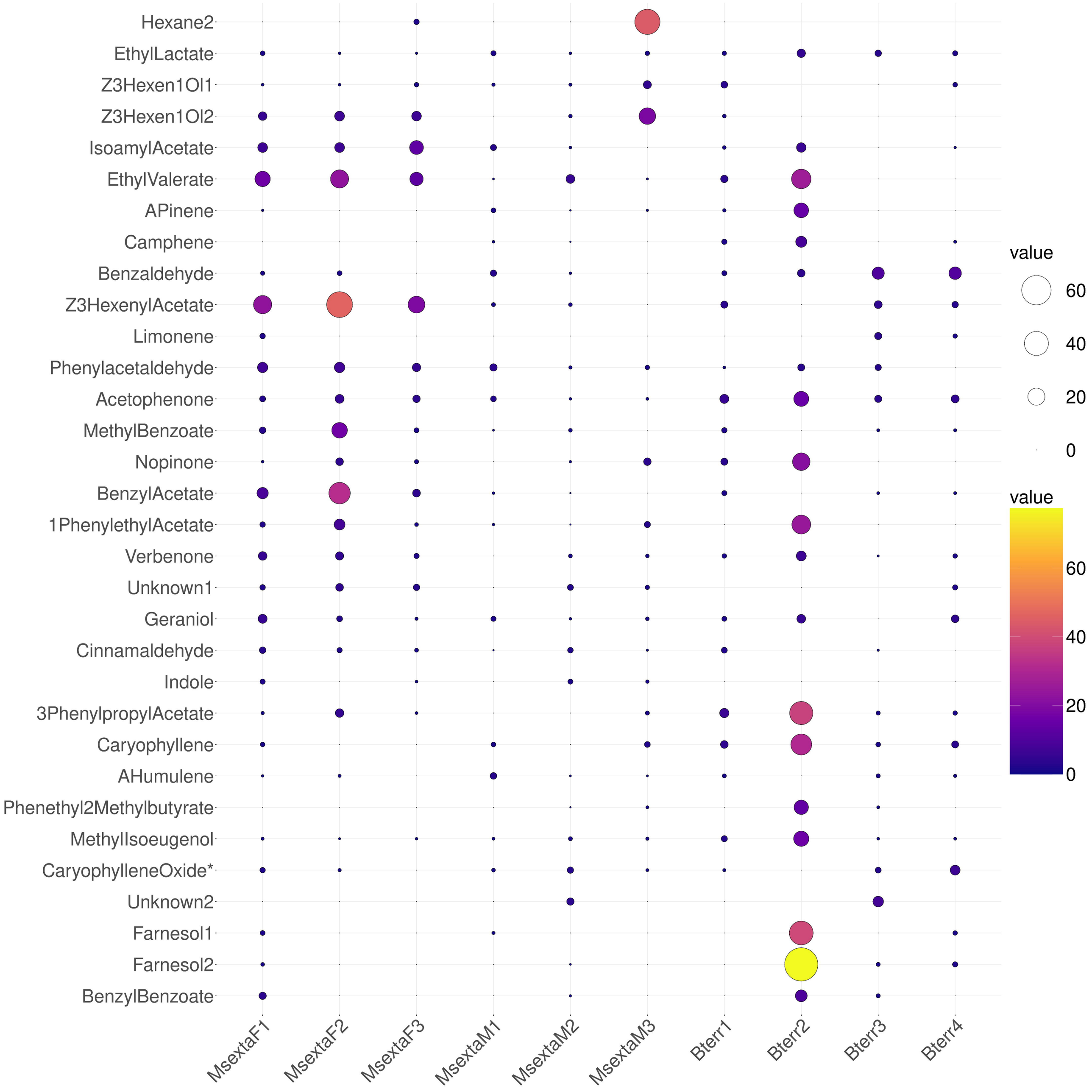
Table S1: Names and addresses of vendors from whom authentic reference standards were sourced.

Table S2: Sources of authentic reference standards used in the EAD test mixture.

Table S3: Literature sources demonstrating previous electroantennogram activity or lack thereof in *Manduca sexta* and *Bombus terrestris* to the compounds used in this study. Note the frequent discrepancies in activity (especially for activity in *B. terrestris*), even for compounds with multiple citations. Compounds with at least one active citation in either species are highlighted in bold. Compounds shaded in yellow were EAD-active in more than 50% of *M. sexta* males or females and/or *B. terrestris* worker individuals from this study.


A
GC-FID Signal
FID amplitude (pA)
Z3Hexen1Ol1
Z3Hexen1Ol2
IsoamylAcetate
EthylValerate
APinene
Camphene
Benzaldehyde
Z3HexenylAcetate
Limonene
Phenylacetaldehyde
Time (minutes)
B
Moth 1 (female): raw EAD amplitude
Raw EAD amplitude (mV)
Time (minutes) (adjusted for GC-FID injection time)
C
Moth 1 (female): first forward difference of antennal response
First forward difference
Time (minutes) (adjusted for GC-FID injection time)


A
Manduca sexta females: EAD significant peak frequency by compound
FAD (n = 6)
Terpenoid (n = 11)
Aromatic (n = 11)
Nitrogenous (n = 1)
Unknown (n = 2)
Solvent (n = 2)
B
Manduca sexta males: EAD significant peak frequency by compound
FAD (n = 6)
Terpenoid (n = 11)
Aromatic (n = 11)
Nitrogenous (n = 1)
Unknown (n = 2)
Solvent (n = 2)
C
Bombus terrestris workers: EAD significant peak frequency by compound
FAD (n = 6)
Terpenoid (n = 11)
Aromatic (n = 11)
Nitrogenous (n = 1)
Unknown (n = 2)
Solvent (n = 2)


Hexane2
EthylLactate
Z3Hexen1Ol1
Z3Hexen1Ol2
IsoamylAcetate
EthylValerate
APinene
Camphene
Benzaldehyde
Z3HexenylAcetate
Limonene
Phenylacetaldehyde
Acetophenone
MethylBenzoate
Nopinone
BenzylAcetate
1PhenylethylAcetate
Verbenone
Unknown1
Geraniol
Cinnamaldehyde
Indole
3PhenylpropylAcetate
Caryophyllene
AHumulene
Phenethyl2Methylbutyrate
MethylIsoeugenol
CaryophylleneOxide*
Unknown2
Farnesol1
Farnesol2
BenzylBenzoate
MsextaF1
MsextaF2
MsextaF3
MsextaM1
MsextaM2
MsextaM3
Bterr1
Bterr2
Bterr3
Bterr4
value
60
40
20
0
value
60
40
20
0