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: 5080

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. Our method also works with existing, published GC-EAD datasets.

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

33 **INTRODUCTION**

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

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

57 Electroantennography – the detection and measurement of electrical signals from whole insect
58 antennae – is often used to screen plant volatiles for insect perceptibility. In this technique, a pair of
59 electrodes are mounted at either end of an excised insect antenna or at the tip of the antenna and
60 another reference point such as the base of the head (typically for smaller insects). The continuous
61 voltage difference across the antenna is then amplified and sent to a computer. The insect antenna can
62 be stimulated using single air puffs passed across filter paper spotted with single compounds or a mix of
63 compounds (electroantennography, abbreviated EAG), or instead can be exposed to the effluent from a
64 gas chromatography (GC) column (GC-coupled electroantennographic detection, or GC-EAD). The two
65 signals – from the GC's main detector (usually a flame ionization detector, or FID) and from the insect
66 antenna – are then aligned and compared by eye to determine whether a response has occurred to a
67 given GC-FID peak. The compounds to which the insect antenna "responds" with some voltage change
68 are taken to be physiologically active and may be further tested via behavioural assays with single pure
69 or mixed compounds or with plant lines known to be high or low in the relevant compounds. While GC-
70 EAD has significant advantages over EAG assays due to its ability to test responses to commercially or
71 synthetically unavailable or unknown compounds present in a chemical mixture, the resulting signal is
72 typically fairly noisy and can be difficult to interpret.

73 Although these techniques have been in use for nearly seventy years, being first described in 1957
74 (Schneider, 1957), the resulting data are always presented qualitatively – "response/no response" (e.g.
75 Peakall et al., 2010) – with no quantitative, repeatable measure used to define antennal "response"
76 versus background activity. Indeed, a comprehensive and well-cited overview of GC-EAD includes no
77 mention of how data analysis is performed or what criteria are used to define an "active" GC-EAD peak
78 (Schiestl and Marion-Poll, 2002). This has hampered the field by limiting reproducibility and cross-study
79 or cross-observer comparisons, introducing error, and reducing ease of application to broader plant
80 volatile studies. This may be a particular issue with many plant volatiles, which – unlike insect

81 pheromones – may not have the same degree of "dedicated hardware" on the insect antenna (Akers
82 and Getz, 1993; Schiestl and Marion-Poll, 2002) and thus may produce weaker signals, with the
83 exception of sexually deceptive insect pollination systems which use insect pheromone compounds to
84 attract pollinators.

85 In this Protocol Note, we present "best practice" statistical approaches incorporating techniques from
86 signal processing to analyse gas chromatography-coupled electroantennographic detection (GC-EAD)
87 data in a quantitative fashion. Signal processing, an engineering technique where mathematical
88 transformations are used to enhance signals and remove background noise, has been used in medical
89 electrophysiology (Venkatachalam et al., 2011; van Drongelen, 2018). To our knowledge it has never
90 been applied to electroantennographic data specifically, despite this being a type of electrophysiological
91 study. Our method uses signal processing analyses for correcting for baseline drift and electrical noise
92 (specifically, the use of the first forward difference and median filtering), background antennal noise
93 (specifically, comparison with response to GC outputs between GC-FID peaks), and antennal sensitivity
94 loss (via use of windowing) are presented. Median filtering is a technique used to reduce impulsive noise
95 and smooth data; it functions by setting the value at a given timepoint to be the median of a
96 surrounding "window" of data of a given length (Lim, 1990), in our case with the goal of removing
97 random electrical noise often present in electroantennography data. The first forward difference (FFD),
98 another signal processing technique, outputs the step change from the value at time t to the value at
99 time $t+n$, where n is 1 for the "first" forward difference used here (Oppenheim and Schafer, 1975). The
100 FFD is the simplest possible way to remove low frequency components, including the baseline drift
101 present in electroantennography data. This baseline drift, often seen as a slow, low frequency wiggle of
102 the antennal response, can superficially resemble legitimate responses and thus needs to be controlled
103 for (Schiestl and Marion-Poll, 2002). As the antenna can lose responsiveness over time and the signal
104 often becomes noisier as it degrades (Byers et al., 2020), our method "windows" the EAD data into

discrete time chunks and calculates outlier values for each time chunk separately, rather than calculating the outliers for the entire data set at once. By doing so, we avoid assuming that a significant response in e.g. the 12-16 minute time window (when the antenna may have degraded somewhat) must look identical to a significant response in the earlier e.g. 0-4 minute time window when the antenna is perhaps more sensitive.

A sample data set from six *Manduca sexta* hawkmoths (three males and three females), four *Bombus terrestris audax* bumble bees (all workers), and 28 common volatile stimuli of different volatile chemical compound classes (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. We show that quantitative analyses of GC-coupled 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

Please see Appendix 1 for a detailed equipment list and protocol. Perl and R scripts may be found at https://github.com/plantpollinator/GCEADAnalysis/

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

127 oviposit freely on live *Nicotiana benthamiana* plants and males and females were kept together prior to
128 collection. Worker *Bombus terrestris audax* (Harris) (Hymenoptera: Apidae) were sourced from a single
129 Bumblebee New Season Hive (Biobest, purchased from Agralan Ltd, Wiltshire, UK) which were allowed
130 to freely forage in a custom-built acrylic cage with 25% sucrose and ground pollen *ad libitum* (Sevenhills
131 Wholefoods Organic Raw Bee Pollen, Dinnington, UK). Both species of insect were maintained prior to
132 sampling in growth chambers at 25°C and constant day length (16 hours light, 8 hours dark).

133 *EAD test mixture preparation and analysis*

134 To assess antennal responses to multiple compounds delivered at approximately equal flame ionization
135 detector (FID) sensitivities, a mixture of 28 authentic reference standards was assembled, representing
136 29 GC-FID peaks (farnesol is sold as a mixture of isomers and both elute separately on the GC-FID).
137 Standards were obtained from a variety of established sources (see Table S1 for vendor addresses and
138 Table S2 for vendor and product numbers) and diluted to 10µg/µl in hexane (Table S2, HPLC-grade, ca.
139 95% purity). The final compound list was selected to include a variety of chemical compound classes
140 (terpenoids, aromatics, fatty acid derived compounds, and one nitrogenous compound) that would
141 separate well on a 10°C per minute GC oven ramp program. Due to differences in FID responsiveness to
142 the individual compounds (i.e. some compounds may produce different peak areas on the FID despite
143 being at equal injected amounts), different volumes of the 10µg/µl dilutions were added to the
144 experimental mixture (see Table S2 for final concentrations, which ranged from 93ng/µl for (R)-(+)-
145 limonene to 2004ng/µl (2.004µg/µl) for indole).

146 To determine retention times and confirm mixture GC-FID peak identities, 3ul of the test mixture was
147 manually injected into an Agilent 7890B GC (Agilent, Stockport, UK) fitted with an Agilent 5977A/B MS
148 detector (GC-MS) and a Phenomenex Zebron ZB-5HT column (35m x 0.25mm x 0.10µm; Phenomenex UK
149 Ltd, Macclesfield, UK). The GC run parameters were as follows: the column flow was 1.0603 ml/min of

150 helium carrier gas with an oven program of 50°C for 4 minutes followed by a temperature ramp of
151 10°C/minute to 230°C, which was then held for 4 minutes. The inlet was splitless and held at 250°C and
152 the MS source was held at 230°C. Compounds were identified using Agilent Unknowns software (v10.1,
153 Agilent, Stockport, UK) and via comparison to Kovats Retention Indices of the individual compounds run
154 separately using the same GC-MS system, calculated using an alkane ladder (49452-U, Sigma-Aldrich,
155 Watford, UK). Two additional "unknown" GC-FID peaks were identified in the final mixture which did not
156 represent any of the chemical standards added or the solvent. As many of the standards used are
157 derived from natural sources rather than synthetic ones and are not completely pure as a result, we
158 expect that these represent additional minor natural products from these sources. Due to our inability
159 to remove these unknown compounds from the mixture, we treated them as a separate chemical
160 compound class and excluded them from downstream analyses that targeted chemical compound
161 classes or modification types.

162 *EAD antennal preparation and assembly*

163 *Manduca sexta* male and female adults: single antennae were excised using microdissection scissors
164 from three individuals each of living adult male and female *Manduca sexta*. Due to biosecurity
165 requirements for this non-native potential plant pest species, excised antennae were placed in a 15ml
166 conical tube on wet tissue paper to maintain antennal humidity levels and carried to the
167 electrophysiology setup in a separate building. The fine tip of an antenna (representing ca. 7 segments,
168 cut where the tip abruptly broadens) was removed using a microscalpel under a dissection microscope.
169 Tip removal is commonly done (though is not mandatory) in electroantennographic studies, as it may
170 reduce electrical resistance and thus improve signal (Schoonhoven 1973). The excised and trimmed
171 antenna was then mounted on an antenna electrode fork (Syntech PRG-2, Ockenfels Syntech,
172 Buchenbach, Germany) using electrode gel (Signa Gel, Parker Laboratories, Fairfield, New Jersey, USA),

173 which was then plugged in to the electrode mount (Syntech MP-15 with EAG Combi Probe) and moved
174 into position in front of the GC eluent-carrying air stream approximately 1cm from the air stream outlet.
175 Approximately ten minutes elapsed between antennal excision and mounting on the antenna electrode
176 fork due to the walk between buildings.

177 *Bombus terrestris audax* workers: individuals were removed from the Entomology platform facility in
178 50ml conical tubes, then kept at 4°C for 30 minutes to enable safe antennal removal. A single antenna
179 was then removed using microdissection scissors under red light (with the insect held using a set of
180 forceps) and the tip of the final flagellar segment removed using a microscalpel under a dissection
181 microscope, then treated as above.

182 *GC-EAD setup and recording procedure*

183 The GC-FID setup consisted of an Agilent 8890 GC fitted with a flame ionization detector (FID). The
184 column (a Phenomenex Zebron ZB-5HT column (35m x 0.25mm x 0.10µm), identical to the one used for
185 GC-MS analysis) was fitted to a splitless inlet kept at 250°C. The detector end of the column was fitted to
186 the bottom end of an inert glass Y-splitter (Agilent UltraInert Universal PressFit Y-Splitter , part 5190-
187 6980). The two other arms of the splitter were fitted with deactivated column segments of equal 30cm
188 length (Agilent CP803210, 0.32mm inner diameter), with one end running to the FID (which was kept at
189 300°C) and the other end leaving the GC oven through a port on the side which was fitted with a
190 Syntech EC-03 heated transfer tube kept at 250°C. The GC used hydrogen as a carrier gas, which was
191 generated with a Parker Hydrogen Generator (H2PEM-100, Thames Restek UK Ltd, Saunderton, United
192 Kingdom). Nitrogen gas at a flow of 25ml/min was used as a make-up gas, and the FID was also fed with
193 compressed air at 400ml/min to ensure adequate flame gas supply. The GC run parameters were as
194 follows: the column flow was 4.7 ml/min with an oven program of 50°C for 4 minutes followed by a
195 temperature ramp of 10°C/minute to 230°C, which was then held for 4 minutes. Each manual sample

injection consisted of 3ul of the EAD test mixture, for a total of 0.279-6.012µg of each compound injected. GC-FID chromatogram waveforms were exported from Agilent OpenLab CDS v2.5 software as CSV files consisting of an FID amplitude value (in pA) for each given sampled timepoint.

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

Methods: Data processing and analysis

GC-FID peak identification

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

Data loading and initial processing

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

228 *Median filtering, First Forward Difference (FFD) processing, and significant FFD value identification*

229 To remove random electrical noise from the EAD signal, we first applied a median filter using the
230 runmed() function in base R (specifically as `runmed(EADamplitude,17)`). A window of 17 samples
231 was chosen based on the width of the random electrical spikes seen in the *Acleris* data from Dr. Glenn
232 Svensson (**Figure 5B**). In order to remove the effects of the "wandering baseline" common to GC-EAD,
233 we used the `diff()` function in base R (specifically as
234 `diff(EADamplitude,lag=1,differences=1)`) to calculate the First Forward Difference of this
235 median filtered antennal amplitude . The FFD vector was subsetted into five windows of four minutes
236 each for a total GC-FID run of 20 minutes for the hawkmoth and bumble bee data. For the data from Dr.
237 Svensson, we instead used three minutes each for a total GC-FID run of 21 and 14.5 minutes for the
238 *Acleris* and *Plodia* data respectively. For each window, the average and standard deviation of the FFD
239 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 and termed "significant spikes".

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

We first calculated how many GC-EAD FFD spikes above the 3σ threshold were within each GC-FID peak. For each GC-FID peak identified, the number of significant GC-EAD FFD spikes as defined above within its onset-offset window was determined using the 'between' function in the R package `dplyr` (Wickham et al., 2023), which identified significant spikes that were within a GC-FID peak window. This value was divided by the length of the GC-FID peak in seconds (offset time minus onset time) to get the significant FFD spike frequency of spikes/second. As EAD data can have significant background noise (Schiestl and Marion-Poll, 2002), we defined the significant FFD spike frequency of the space between GC-FID peaks as the threshold for whether a GC-FID peak attracted significant EAD activity for each window: if the significant FFD spike frequency for a given GC-FID peak exceeded the background significant FFD spike frequency of its time window, the GC-FID peak's EAD response was counted as "EAD active", and otherwise considered non-active.

Statistical analysis

To compare results between individuals within a sex/caste and between species, a "significant response" across each species-sex/caste category (female *M. sexta*, male *M. sexta*, and worker *B. terrestris*) was identified for chemical compounds where >2/3 of individuals tested were "EAD active" for that peak. To investigate whether the results were repeatable between moths, we calculated the Pearson correlation coefficient between the significant spike frequency for all pairwise combinations of moths within a sex using cor.test() in R. We then used the linear model function `lm()` in R to calculate the slope and intercept of the relationship between each pair of moths within a sex. Finally, we used `lm()` and the estimated marginal means function `emmeans()` (Lenth, 2024) to determine whether

263 chemical compound class (e.g. terpenoid) or modification type (e.g. alcohol) had an influence on
264 significant spike frequency within each species-sex/caste group (female *M. sexta*, male *M. sexta*, and
265 worker *B. terrestris*).

266 *Further test with unpublished and published data*

267 Unpublished and published GC-EAD data were kindly donated by Dr. Glenn Svensson (Pheromone
268 Group, Lund University). Data from *Plodia interpunctella* (Lepidoptera: Pyralidae) were of a single male
269 in response to a synthetic mixture of four female pheromone components. Data from *Acleris comariana*
270 (Lepidoptera: Tortricidae) were of a single male in response to a natural female pheromone gland
271 extract. As we had difficulty employing our GC-FID automated peak detector on these data due to
272 significant noise and a long solvent peak trailing end in the GC-FID signal, we instead manually
273 annotated GC-FID peaks in GcEad 2014 software and used those retention times in the downstream
274 analysis. Due to the different length of the GC-EAD runs, we chose 3 minute windows for calculation of
275 significant FFD spikes instead of 4 minute windows as in our earlier experiments.

276 **Results**

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

278 Here we present a method that avoids the otherwise significant biases inherent in GC-EAD *qualitative*
279 analysis methodology by the use of signal processing techniques, specifically *median filtering* and the
280 *first forward difference* (FFD), to remove the influence of electrical noise and baseline drift respectively,
281 essentially flattening out the EAD signal (**Figure 1B**) and replacing its waveform with a series of
282 individual spikes (**Figure 1C**, see also **Figures S16-S30**). A significance threshold may then be set for each
283 window (horizontal lines in **Figure 1C** and **Figures S16-S30**, both showing ± 3 standard deviations
284 above/below the mean for that window), and the frequency of spikes with an amplitude over that

285 significance threshold during a given GC-FID peak may be easily calculated and analysed using traditional
286 statistical methods. This results in a *quantitative* rather than *qualitative* assessment of whether a GC-FID
287 peak is physiologically active, i.e. provokes a significant GC-EAD response, or not.

288 *GC-FID peaks with significant EAD responses*

289 **Table 1** shows the list of compounds tested in the EAD test mixture along with their significant spike
290 rates for female and male moths and worker bumble bees and how many individuals of each had a
291 significant spike rate over the background spike rate. In general, female moths were most responsive to
292 fatty acid-derived compounds (FADs, red in **Figure 2A**, see also **Figure S43**) and to some but not all
293 aromatic volatiles (blue in **Figure 2A**). Responses to terpenoids were more mixed (green in **Figure 2A**),
294 with some compounds (e.g. (geraniol) eliciting significant spike frequencies, while others (e.g.
295 camphene) eliciting no significant spike requencies in any female. The sole tested nitrogenous
296 compound, indole (yellow in **Figure 2A**), also elicited significant spike frequencies in all three female
297 moths. Overall, female moths responded significantly to 20 of the 31 tested compounds (excluding the
298 hexane GC-FID peaks). Male moths were generally less responsive than females to the majority of
299 compounds, with overall significant spike frequencies to 20 of the 31 tested compounds (excluding the
300 hexane GC-FID peaks) but with no significant difference in significant spike frequency between chemical
301 compound classes (e.g. terpenoids versus aromatics) (**Figure S43A,B**); 2 of 3 males were responsive to
302 indole. Worker bumble bees responded significantly to 8 of the 30 compounds ("unknown 2" was
303 missing from this analysis as it was not picked up by the GC-FID peak detector script). Bees responded
304 equally (i.e. had equal significant spike frequencies) to all types of chemical compound classes (**Figure**
305 **S43C**), although only one of four tested bees was significantly responsive to indole, the sole nitrogenous
306 compound tested.

307 Compounds containing oxygenated modifications, particularly esters, seemed to produce the highest
308 level of significant spike frequency in female moths, but not in males (**Figure S43D,E**). At least two-thirds
309 of females responded significantly to 8 of 10 presented esters (analyzed without the co-eluting GC-FID
310 peaks ethyl laurate and caryophyllene oxide), 3/6 alcohols, 3/3 aldehydes, and 3/3 ketones, but not to
311 any of the five presented unmodified monoterpenes ((1R)-α-pinene, camphene, and R-(+)-limonene) or
312 sesquiterpenes (caryophyllene, α-humulene), both groups of which lack oxygen atoms. Male moths
313 responded to 6 of 10 esters, 3/6 alcohols, 2/3 aldehydes, and 2/3 ketones but, unlike females,
314 responded to 3/5 of the unmodified (mono/sesqui)terpenes. Males also responded significantly overall
315 to the "Unknown 2" compound, whose identity remains unclear despite GC-MS analysis but which likely
316 also is oxygenated based on tentative NIST library matches. For bees, there was no difference in
317 response based on compound modification, oxygenated or otherwise (**Figure S43F**). At least 2/3 of bees
318 responded to 5 of 10 esters, 1/3 aldehydes, and 2/5 unmodified (mono/sesqui)terpenes, but fewer than
319 2/3 of bees responded to the 6 alcohols or 3 ketones.

320 *Repeatability within moth sexes and comparisons between sexes*

321 Responses to non-hexane compounds were highly repeatable between female moths (**Figure S44A**,
322 Pearson's correlation coefficient range 0.537 to 0.762, all significant), but not between male moths
323 (**Figure S44B**, Pearson's correlation coefficient range 0.025 to 0.476, only one significant). In general,
324 male moths were substantially less responsive to non-hexane chemical stimuli in this experiment, as
325 represented by their lower average significant spike frequency (**Table 1,** range 0.051-8.333 significant
326 spikes/second) versus females (range 0.159-12.222 significant spikes/second). Responses between
327 males and females trended towards correlation (Pearson's r = 0.342, $P = 0.051$, **Figure S45**).

328

329 *Validation of the method with existing external data*

330 To validate our method with existing GC-EAD data from another lab and setup, we obtained both
331 published and unpublished GC-EAD data from Dr. Glenn Svennson (Pheromone Research Group, Lund
332 University). These data represent responses of *Plodia interpunctella* (Lepidoptera: Pyralidae; **Figure 4,**
333 **Table S3**) and *Acleris comariana* (Lepidoptera: Tortricidae; **Figure 5, Table S4**) to synthetic and natural
334 pheromones respectively. In both cases our method was able to identify significant GC-EAD peaks,
335 including confirming the published *A. comariana* response to the pheromone component (*E*)-11,13-
336 tetradecadienal (Svensson et al., 2019). Of note, we identified multiple significant spike frequency
337 responses in both datasets that were not identified as significant by previous qualitative analyses
338 (**Tables S3, S4,** Svensson et al., 2019); as our method more unambiguously identifies low-amplitude
339 responses, we suggest ranking potential compounds of interest by their significant FFD hit frequency
340 when prioritizing which compounds to take further via behavioural analysis. The most significant spike
341 frequency by far in the *A. comariana* data, for example, was to the published bioactive compound
342 (**Figure 5, Table S4**).

343

344 **CONCLUSIONS**

345 In this Protocol Note, we have presented a novel signal processing-based methodology for use with gas
346 chromatography-coupled electroantennographic detection (GC-EAD) measurements of insect antennae
347 in response to plant volatiles. Both female and male *Manduca sexta* hawkmoths were responsive to
348 approximately 2/3 of tested compounds in a highly repeatable fashion, although male hawkmoths of the
349 same species were far less repeatable between individuals. In female moths, both chemical compound
350 class (FAD, aromatic, terpenoid, or nitrogenous) and modification (none, alcohol, aldehyde, ester,
351 ketone, or oxide) played a significant role in response patterns, while in male moths this was true for
352 neither modification type nor chemical compound class. Worker *Bombus terrestris audax* bumble bees

showed an intermediate level of response, with significant spike frequency responses to one quarter of tested compounds, and with no difference in response rate relative to chemical compound class or modification.

To ensure that our results were not simply due to some property of *Manduca sexta* antennae, we also analysed the same stimulus mixture with four workers of the polylectic (wide dietary breadth, Rasmont et al., 2008) bumble bee *Bombus terrestris audax* using the same data collection and analysis methodologies. We found three compounds that were significant in >2/3 of tested individuals across all three species-sex/caste combinations (female hawkmoths, male hawkmoths, worker bumble bees): ethyl L-lactate, (*Z*)-3-hexenyl acetate, and methyl benzoate (**Table 1**). Results from all three groups are presented in **Figure 3** and **Table 1**. We also observed overall lower significant spike frequencies in bumble bees compared with hawkmoths, despite bumble bees' robust responses to single puffs of the EAD stimulus mixture. The signal to noise ratio threshold in the bumble bee antennae appears to be weaker than in the hawkmoths (as shown by much lower first forward difference significance thresholds and a lower range of significant average spike frequencies (range 0.594-4.715 hits/second compared with hawkmoth values of 0.159-12.222 and 0.594-4.715 for females and males respectively). We separately conducted a non-exhaustive literature review to identify whether our tested compounds had previously been identified as electroantennographically active in either *Manduca sexta* or *Bombus terrestris*, with the results shown in **Table S3**. In general, many of our tested compounds had not previously been tested with GC-EAD or puff-based (without GC separation) electroantennography, and those that had been often produced results that differed between studies, further suggesting the need for a quantitative analysis technique such as the one presented here.

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

376 relying on one or two individuals, given that we found variability even between individuals of the same
377 commercial bumble bee hive and hawkmoth colony. The variability commonly seen in GC-EAD
378 experiments – including ours - can also be due to experimental factors such as humidity, gel coverage of
379 antennal sensillae, insect age and time between antennal excision and mounting. The weaker responses
380 seen in bumble bees further necessitate the use of this new technique to ensure unbiased assessment
381 of electroantennographic detection studies in potentially less strongly responsive insects.

382 By studying responses to common plant volatiles in two model pollinator species, we have shown that
383 responses can be species- and sex-specific, and that it may be difficult to generalize results from one
384 pollinator group to another, or even between populations or genotypes of the same species. Further
385 study will be needed to determine whether results can be generalized within a genus of pollinators, or
386 between those that are generalists and specialists. However, at least one species of pollinating insect
387 was responsive to the majority of the plant compounds tested here. Interestingly, there was no
388 consistent GC-EAD response in any species tested to the monoterpene (R)-(+)-limonene. Limonene is the
389 most widely distributed floral volatile in terms of angiosperm families in which it is found (Schiestl,
390 2010), and is known to provoke a responsive in the antennal lobe of *Bombus vosnesenskii* (Byers et al.,
391 2014a).

392 The data analysis methods introduced make use of existing typical GC-EAD recording setups and
393 experimental methods, differing only in the downstream processing of collected GC-EAD data, and thus
394 may be used with both existing as well as novel datasets. In addition, they are simple to execute,
395 requiring only minimal implementation in R or another statistical processing language (example code is
396 provided in the Appendix). This study provides the first truly quantitative analysis of insect responses to
397 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 expertise to the data analysis methodology development. Both authors approved the final version of the manuscript.

ACKNOWLEDGMENTS

The authors thank the John Innes Centre (JIC) Entomology platform for maintaining the *Manduca sexta* and *Bombus terrestris* colonies and the JIC Metabolomics platform for GC-MS maintenance. We also thank Dr. Glenn Svensson (Pheromone Research Group, Lund University) for providing published and unpublished GC-EAD data, allowing us to test our method on well-validated results from another GC-EAD equipment setup. Funding was provided by the UK Biotechnology and Biological Sciences Research Council Institute Strategic Programmes (Harnessing Biosynthesis for Sustainable Food and Health (HBio) (grant no. BB/X01097X/1) and Building Robustness in Crops (BRiC) (grant no. BB/X01102X/1)). We also thank anonymous reviewers and the editor for their helpful suggestions, which greatly enhanced the manuscript and method.

DATA AVAILABILITY

All raw data (GC-FID CSV files, electroantennography CSV files, and relevant GC-FID injection timestamps) are available on FigShare (doi: XXX). Scripts referenced in the main text and appendices are also available on GitHub (https://github.com/plantpollinator/GCEADAnalysis).

REFERENCES

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

Byers, K. J. R. P. 2021. "As if they discovered it by the scent": improving our understanding of the chemical ecology, evolution, and genetics of floral scent and its role in pollination. *American Journal of 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].

Lim, J. S. 1990. Two-dimensional Signal and Image Processing. Prentice Hall, New Jersey, USA.

Oppeneim, A. V., and R. W. Schafer. 1975. Digital Signal Processing. Prentice-Hall, New Jersey, USA.

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.

Schoonhoven, L. M. 1973. Electroantennograms (EAG) as a tool in the analysis of insect attractants. In P. Grijpma [ed.], Proceedings of the First Symposium on Integrated Control of *Hypsipyla*, Turrialba, Costa Rica, 5-12 March 1973, 59-63.

Svensson, G. P., V. Toennberg, and L. Sigsgaard. 2019. Identification and field evaluation of (*E*)-11,13-tetradecadienal as sex pheromone of the strawberry tortrix (*Acleris comariana*). *Journal of Applied Entomology* 143: 535-541.

van Drongelen, W. 2018. Signal Processing for Neuroscientists, Second Edition. Elsevier, London, United Kingdom.

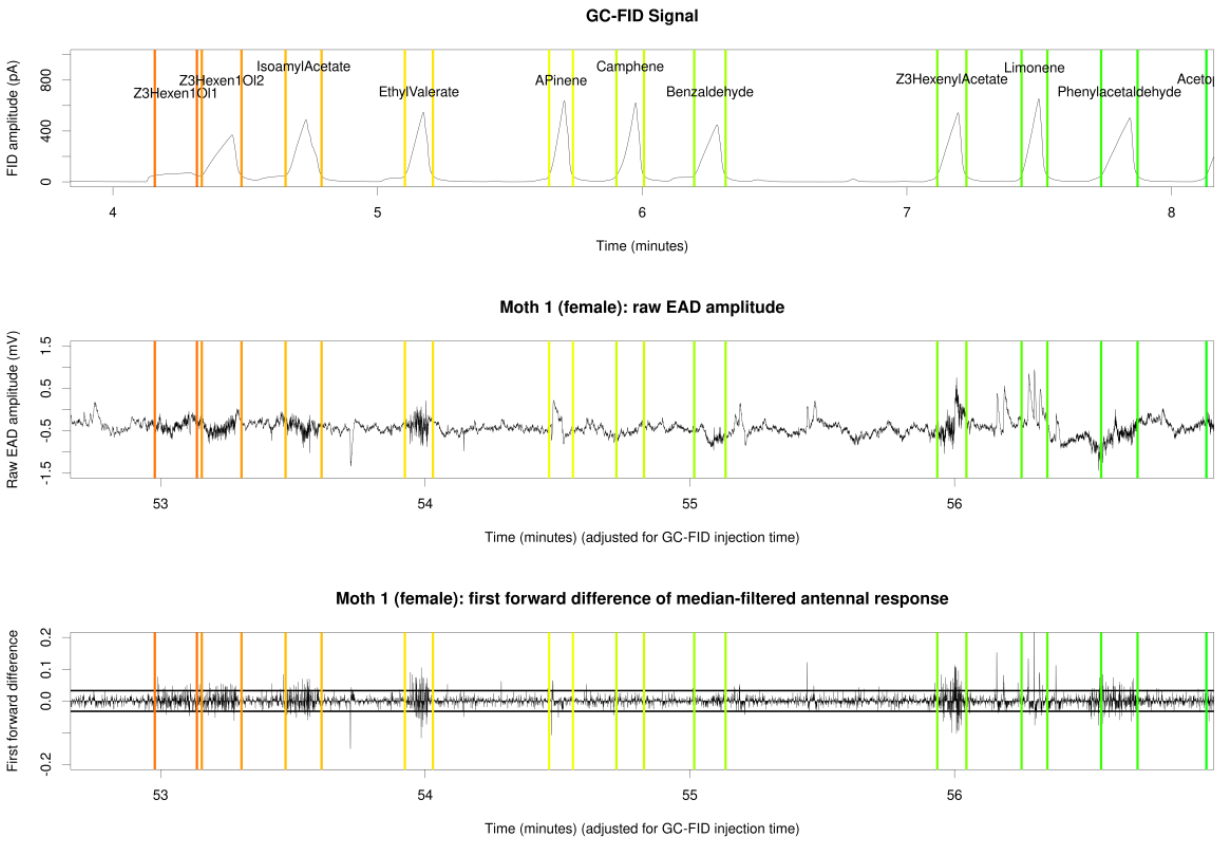
Venkatachalam, K. L., J. E. Herbrandson, and S. J. Asirvatham. 2011. Signals and signal processing for the electrophysiologist: Part II: Signal processing and artifact. *Circulation: Arrhythmia and Electrophysiology* 4: 974-981.

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

Appendix 1 contains a Protocol for collection and analysis of GC-EAD data. Cited Perl and R scripts from the Methods may be found at https://github.com/plantpollinator/GCEADAnalysis/

499 **FIGURES**

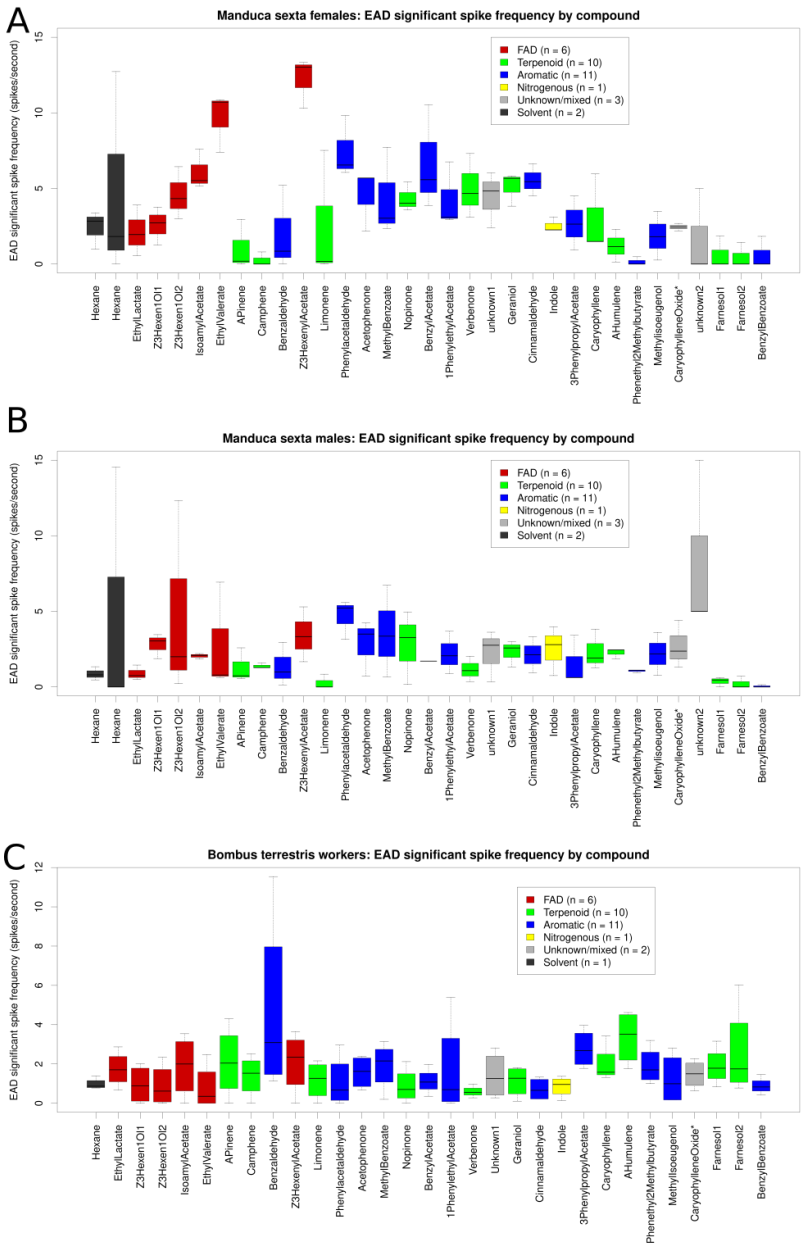


500

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

507

508

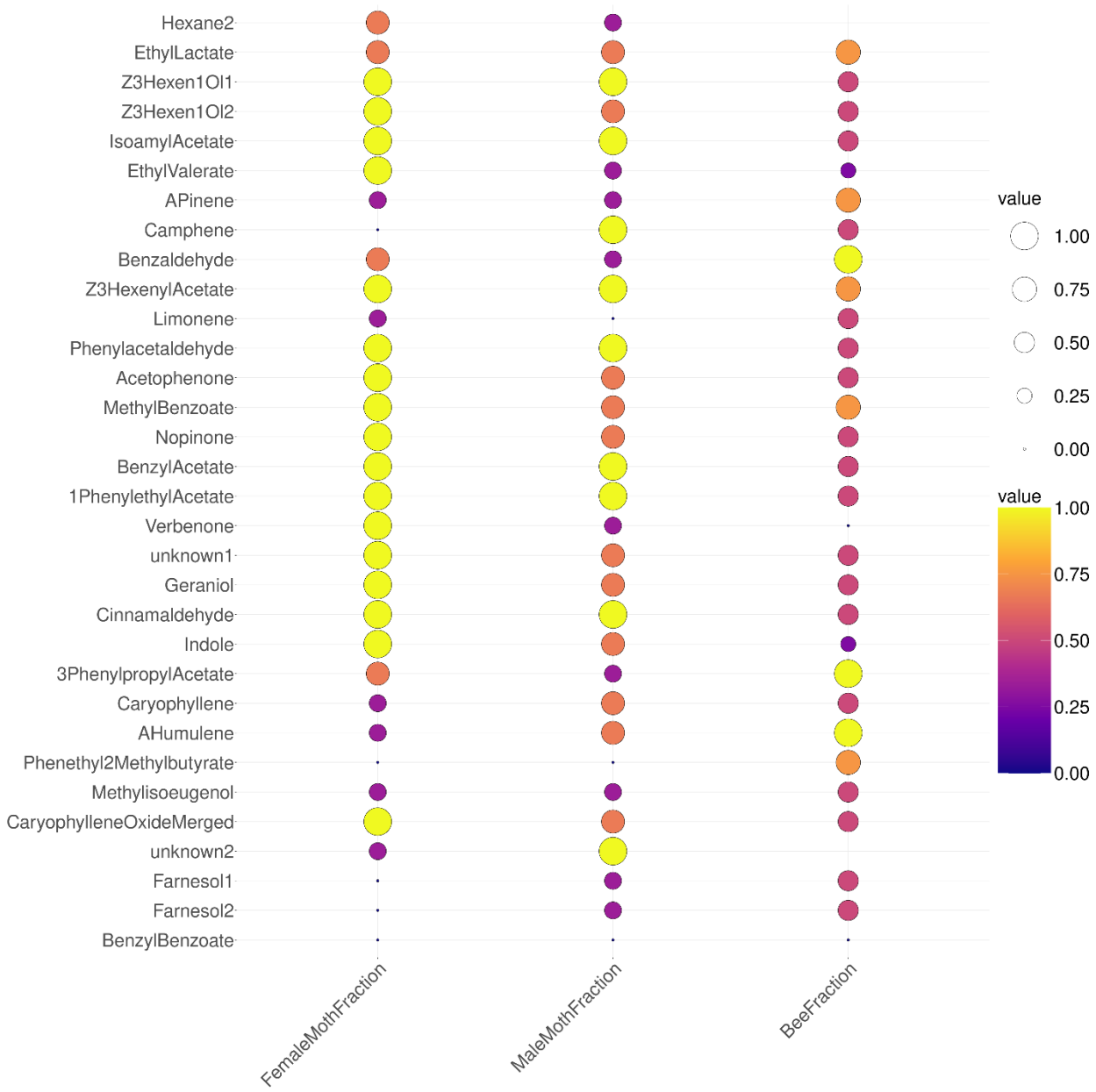


509

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

516

517



518

519 **Figure 3:** Proportion of individual *Manduca sexta* hawkmoths ("FemaleMothFraction" and
520 "MaleMothFraction") and *Bombus terrestris audax* worker bumble bees ("BeeFraction") that had
521 significant responses to each chemical compound tested. Larger circles and warmer colors indicate that
522 more of the individuals were responsive to each compound, while smaller circles and cooler colors
523 indicate that no individuals or only one were responsive. Absent points indicate the specific peak was
524 not detected in the GC-FID run for a given insect.

525

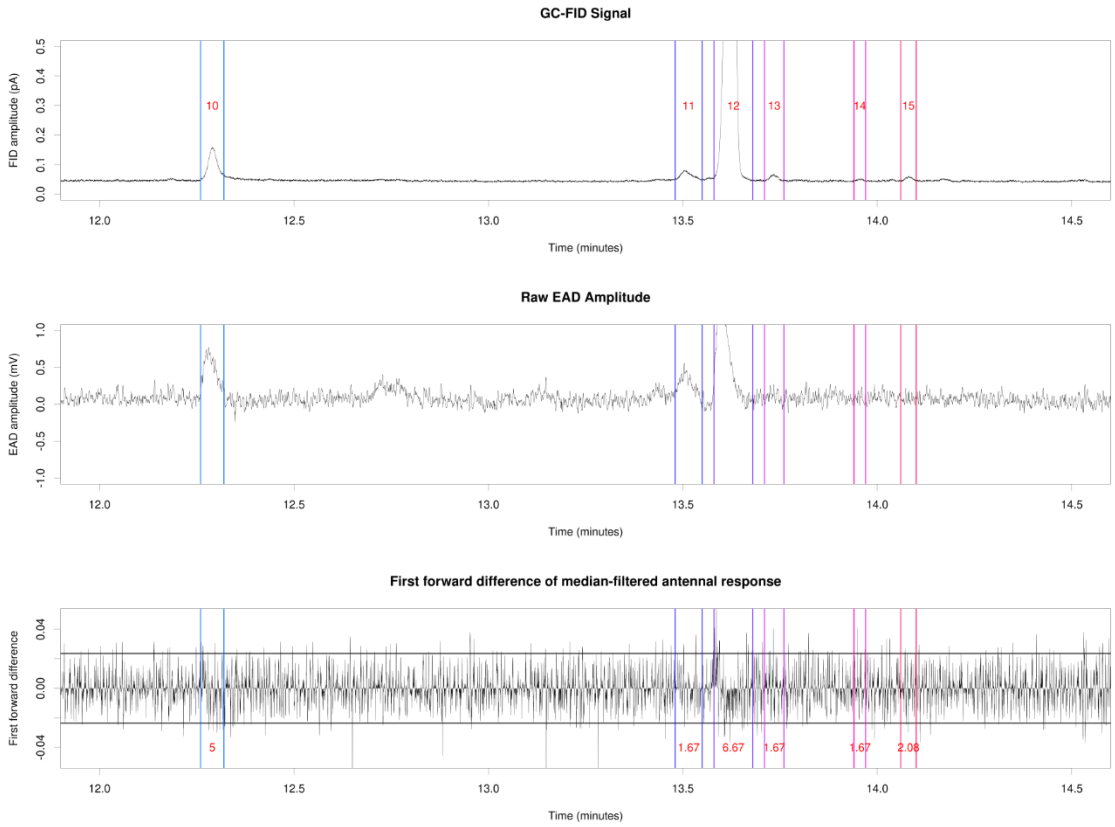


Figure 4: Reanalysis of unpublished GC-EAD data from the Pheromone Group, University of Lund (raw data kindly provided by Dr. Glenn Svensson). Data show the response of a male *Plodia interpunctella* (Lepidoptera: Pyralidae) to a synthetic mixture of four female pheromone compounds. (A) GC-FID trace annotated with manually identified GC-FID peak starts and ends (colored vertical lines) and GC-FID peak numbers. (B) Raw EAD amplitude annotated with GC-FID peak starts and ends. (C) First forward difference-processed median-filtered EAD amplitude annotated with GC-FID peak starts and ends. GC-FID peaks showing significant EAD spike frequency are annotated with that frequency in red text in (C) (all annotated GC-FID peaks within this time window).

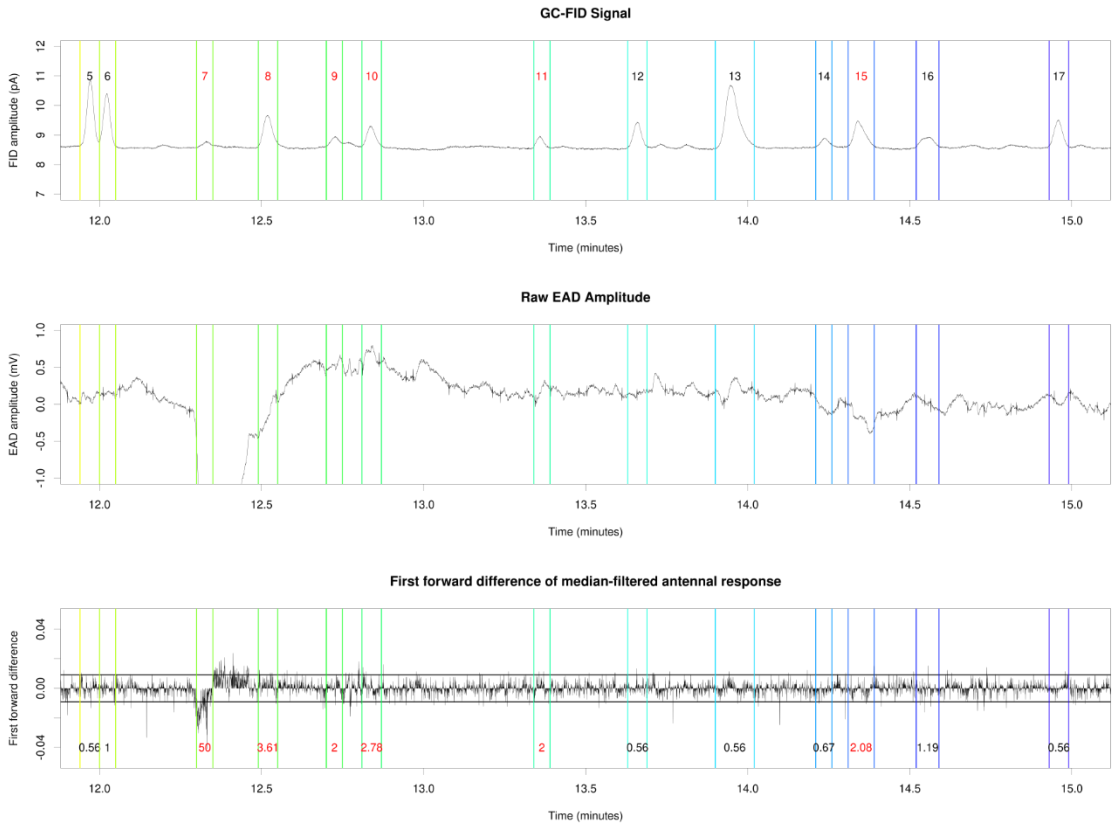


Figure 5: Reanalysis of published GC-EAD data from the Pheromone Group, University of Lund (raw data kindly provided by Dr. Glenn Svensson) (Svensson et al., 2019). Data show the response of a male *Acleris comariana* (Lepidoptera: Tortricidae) to an extract of the female pheromone gland. (A) GC-FID trace annotated with manually identified GC-FID peak starts and ends (colored vertical lines) and GC-FID peak numbers. (B) Raw EAD amplitude annotated with GC-FID peak starts and ends; note the regular small spikes every few seconds, presumably due to electrical noise, which were later removed using a median filter. (C) First forward difference-processed median-filtered EAD amplitude annotated with GC-FID peak starts and ends. GC-FID peaks showing significant EAD spike frequency are annotated with that frequency in red text in (C). GC-FID peak 7 corresponds to (*E*)-11,13-tetradecadienal, the biologically active pheromone component in this species. Note that other GC-FID peaks also provoked a significant response in the male antenna, although they are not listed as EAD-active in (Svensson et al., 2019).

549 **TABLES**

550 **Table 1:** Compounds used in the EAD test mixture with details of their retention time, chemical
551 compound class and subclass or modification categorization, and average significant spike frequency
552 (spikes per second) for female and male *Manduca sexta* hawkmoths and worker *Bombus terrestris*
553 *audax* bumble bees. The ratio after the significant spike frequency indicates the number of samples
554 where the given compound elicited a greater response than the background level of spikes versus the
555 total number of samples with that peak present. Bolded responses: >2/3 of antennal samples reacted to
556 that peak for a given species-sex/caste combination. Unknown superscripts: 1^{A}: an unclear oxygenated
557 cyclic molecule from GC-MS results; 2^{B}: likely an unknown sesquiterpenoid oxide from GC-MS analysis
558 results.

Retention time (GC-FID, minutes)	**Compound identity**	**Chemical compound class**	**Modifications / subtype**	**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 GC-FID peak	Solvent	NA	2.408 (NA)	0.873 (NA)	0.975 (NA)
2.687	Hexane, small GC-FID peak	Solvent	NA	4.848 (NA)	4.848 (NA)	NA (NA)
3.755	Ethyl L-lactate	FAD	Ester	**2.133 (2/3)**	**0.900 (2/3)**	**1.738 (3/4)**
4.291	(*Z*)-3-hexen-1-ol, shoulder GC-FID peak	FAD	Alcohol	**2.583 (3/3)**	**2.792 (3/3)**	0.950 (2/4)
4.458	(*Z*)-3-hexen-1-ol, main GC-FID peak	FAD	Alcohol	**4.593 (3/3)**	**4.852 (2/3)**	0.897 (2/4)
4.741	Isoamyl acetate	FAD	Ester	**6.094 (3/3)**	**2.045 (3/3)**	1.886 (2/4)
5.184	Ethyl valerate	FAD	Ester	**9.659 (3/3)**	2.782 (1/3)	0.796 (1/4)
5.716	(1R)-(+)-α-pinene	Terpenoid	Monoterpene	1.049 (1/3)	1.296 (1/3)	**2.097 (3/4)**
5.985	Camphene	Terpenoid	Monoterpene	0.267 (0/3)	**1.387 (3/3)**	1.396 (2/4)
6.292	Benzaldehyde	Aromatic	Aldehyde	**2.019 (2/3)**	1.362 (1/3)	**4.715 (4/4)**
7.200	(*Z*)-3-hexenyl acetate	FAD	Ester	**12.222 (3/3)**	**3.434 (3/3)**	**2.087 (3/4)**
7.505	(R)-(+)-limonene	Terpenoid	Monoterpene	2.564 (1/3)	0.285 (0/3)	1.176 (2/4)
7.847	Phenylacetaldehyde	Aromatic	Aldehyde	**7.475 (3/3)**	**4.646 (3/3)**	1.074 (2/4)

8.222	Acetophenone	Aromatic	Ketone	**4.526 (3/3)**	**2.822 (2/3)**	1.581 (2/4)
8.691	Methyl benzoate	Aromatic	Ester	**4.370 (3/3)**	**3.585 (2/3)**	**1.912 (3/4)**
9.355	(1R)-(+)-nopinone	Terpenoid	Monoterpenoid ketone	**4.341 (3/3)**	**2.791 (2/3)**	0.885 (2/4)
9.818	Benzyl acetate	Aromatic	Ester	**6.667 (3/3)**	**1.705 (3/3)**	1.126 (2/4)
10.279	1-phenylethyl acetate	Aromatic	Ester	**4.265 (3/3)**	**2.206 (3/3)**	1.696 (2/4)
10.511	(1S)-(-)-verbenone	Terpenoid	Monoterpenoid ketone	**5.030 (3/3)**	1.158 (1/3)	0.594 (0/4)
10.609	Unknown 1[A]	Unclear	Unclear	**4.425 (3/3)**	**2.241 (2/3)**	1.408 (2/4)
11.240	Geraniol	Terpenoid	Monoterpenoid alcohol	**5.111 (3/3)**	**2.306 (2/3)**	1.117 (2/4)
11.459	(*E*)-cinnamaldehyde	Aromatic	Aldehyde	**5.532 (3/3)**	**2.128 (3/3)**	0.718 (2/4)
11.777	Indole	Nitrogenous	Aromatic	**2.545 (3/3)**	**2.509 (2/3)**	0.856 (1/4)
12.691	3-phenylpropyl acetate	Aromatic	Ester	**2.687 (2/3)**	1.550 (1/3)	**2.772 (4/4)**
13.254	β-caryophyllene	Terpenoid	Sesquiterpene	2.979 (1/3)	**2.340 (2/3)**	1.974 (2/4)
13.683	α -humulene	Terpenoid	Sesquiterpene	1.199 (1/3)	**2.254 (2/3)**	**3.352 (4/4)**
14.111	Phenethyl 2-methylbutyrate	Aromatic	Ester	0.159 (0/3)	1.058 (0/3)	**1.900 (3/4)**
14.285	Methyl isoeugenol	Aromatic	Alcohol	1.849 (1/3)	2.194 (1/3)	1.240 (2/4)
15.265	Caryophyllene oxide + ethyl laurate (GC-FID peaks merged by GC-FID peak identification script)	Terpenoid + FAD	Sesquiterpenoid oxide + ester	**2.449 (3/3)**	**2.694 (2/3)**	1.477 (2/4)
15.476	Unknown 2[B]	Terpenoid	Unknown sesquiterpenoid oxide	1.667 (1/3)	**8.333 (3/3)**	NA (NA)
16.474	Farnesol, isomer 1	Terpenoid	Sesquiterpenoid alcohol	0.615 (0/3)	0.359 (1/3)	1.895 (2/4)
16.734	Farnesol, isomer 2	Terpenoid	Sesquiterpenoid alcohol	0.476 (0/3)	0.238 (1/3)	2.573 (2/4)
17.189	Benzyl benzoate	Aromatic	Ester	0.611 (0/3)	0.051 (0/3)	0.885 (0/4)

SUPPORTING INFORMATION: SUPPLEMENTAL FIGURE LEGENDS AND TABLE

Figures S1-S5: Raw EAD amplitude for three female hawkmoths. Top pane: GC-FID trace annotated with GC-FID peak starts and ends (colored vertical lines) and GC-FID peak identities. Bottom three panes: individual female hawkmoth EAD traces annotated with GC-FID 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 GC-FID peak starts and ends (colored vertical lines) and GC-FID peak identities. Bottom three panes: individual male hawkmoth EAD traces annotated with GC-FID 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 GC-FID peak starts and ends (colored vertical lines) and GC-FID 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 GC-FID peak starts and ends (colored vertical lines) and GC-FID peak identities. Bottom three panes: individual female hawkmoth first forward difference-processed EAD traces annotated with GC-FID peak starts and ends (colored vertical lines) and significant GC-EAD spike hit frequency (spikes/second), with red indicating frequency over the background 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 GC-FID peak starts and ends (colored vertical lines) and GC-FID peak identities. Bottom three panes: individual male hawkmoth first forward difference-processed EAD traces annotated with GC-FID peak starts and ends (colored vertical lines) and significant GC-EAD spike hit frequency (spikes/second), with red indicating frequency over the background 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 GC-FID peak starts and ends (colored vertical lines) and GC-FID peak identities. Bottom three panes: individual bumble bee worker first forward difference-processed EAD traces annotated with GC-FID peak starts and ends (colored vertical lines) and significant GC-EAD spike

hit frequency (spikes/second), with red indicating frequency over the background 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.

Figures S31-S35: Reanalysis of unpublished GC-EAD data from the Pheromone Group, University of Lund (raw data kindly provided by Dr. Glenn Svensson). Data show the response of a male *Plodia interpunctella* (Lepidoptera: Pyralidae) to a synthetic mixture of four female pheromone compounds. (A) GC-FID trace annotated with manually identified GC-FID peak starts and ends (colored vertical lines) and GC-FID peak numbers. (B) Raw EAD amplitude annotated with GC-FID peak starts and ends. (C) First forward difference-processed median-filtered EAD amplitude annotated with GC-FID peak starts and ends. GC-FID peaks showing significant EAD spike frequency are annotated with that frequency in red text in (C) (all annotated GC-FID peaks within this time window). Figure S31: 0-3 minutes of GC-FID elapsed time. Figure S32: 3-6 minutes. Figure S33: 6-9 minutes. Figure S34: 9-12 minutes. Figure S35: 12-14.5 minutes.

Figures S36-S42: Reanalysis of published GC-EAD data from the Pheromone Group, University of Lund (raw data kindly provided by Dr. Glenn Svensson) (Svensson et al., 2019). Data show the response of a male *Acleris comariana* (Lepidoptera: Tortricidae) to an extract of the female pheromone gland. (A) GC-FID trace annotated with manually identified GC-FID peak starts and ends (colored vertical lines) and GC-FID peak numbers. (B) Raw EAD amplitude annotated with GC-FID peak starts and ends; note the regular small spikes every few seconds, presumably due to electrical noise, which were later removed using a median filter. (C) First forward difference-processed median-filtered EAD amplitude annotated with GC-FID peak starts and ends. GC-FID peaks showing significant EAD spike frequency are annotated with that frequency in red text in (C). Figure S36: 0-3 minutes of GC-FID elapsed time. Figure S37: 3-6 minutes. Figure S38: 6-9 minutes. Figure S39: 9-12 minutes. Figure S40: 12-15 minutes. Figure S41: 15-18 minutes. Figure S42: 18-21 minutes.

Figure S43: First forward difference-processed EAD significant peak hit frequency (spikes/second) across (A,B,C) four chemical compound 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 S44: 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.

Figure S45: Inter-sex EAD response (repeatability) comparisons between the average values of male and female *Manduca sexta* hawkmoths for each GC-FID peak.

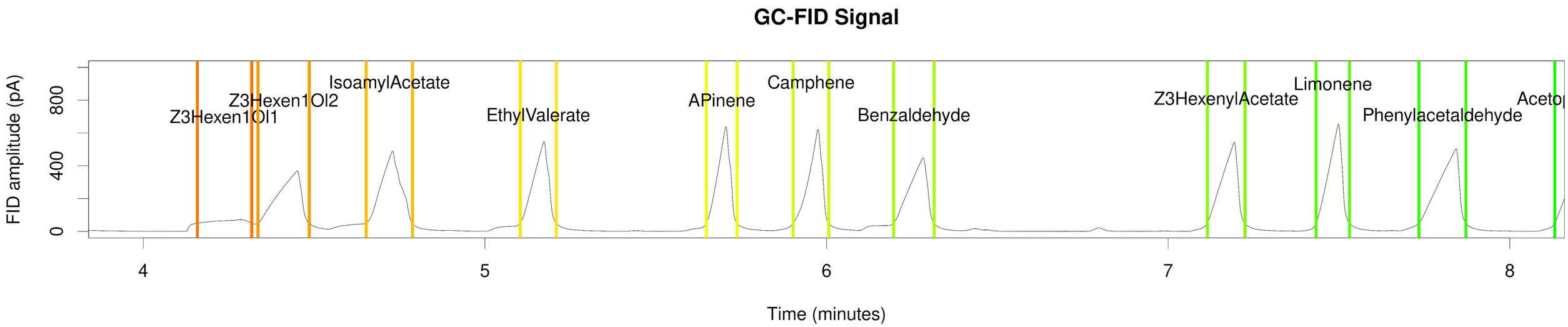
635 **Table S1:** Names and addresses of vendors from whom authentic reference standards were sourced.

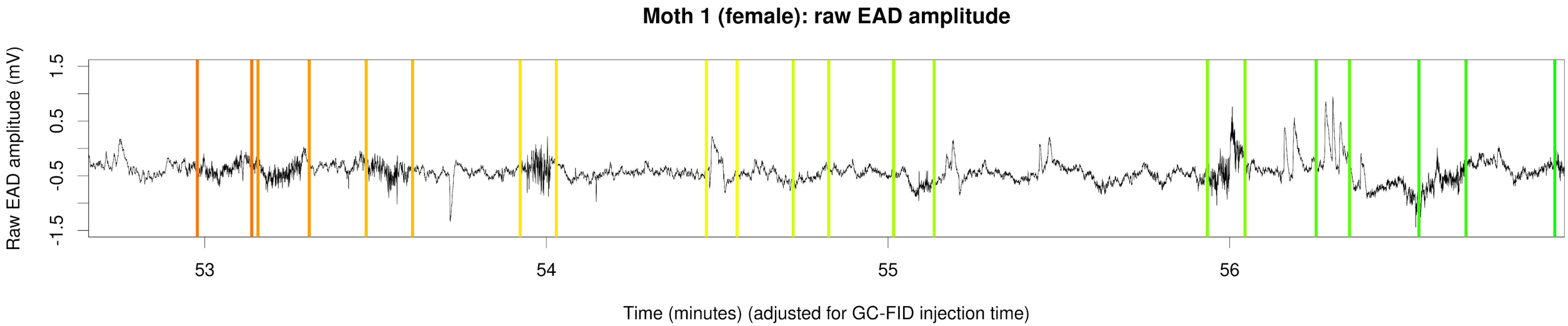
636 **Table S2:** Sources of authentic reference standards used in the EAD test mixture.

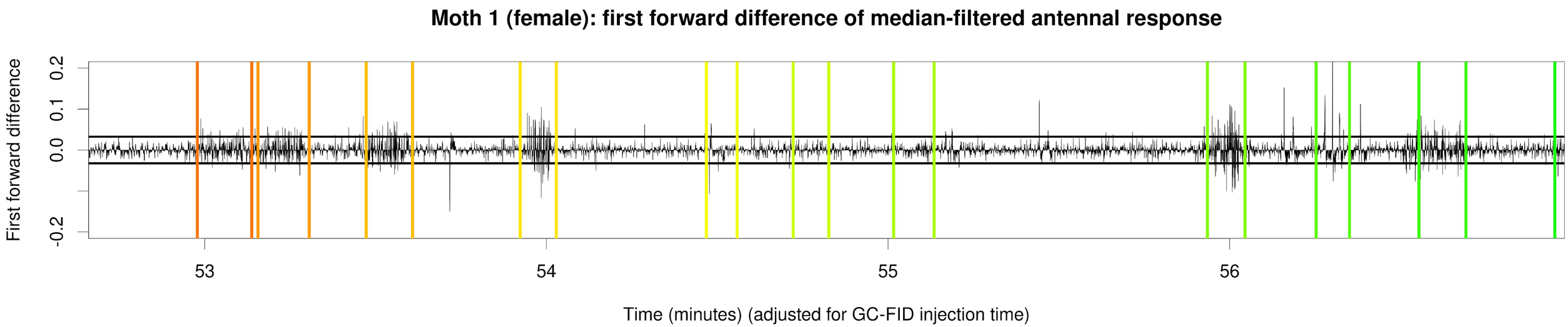
637 **Table S3:** GC-EAD responses by *Plodia interpunctella* to a test mixture of 4 pheromone components.
638 Bolded response values with * were significant above the background response.

639 **Table S4:** GC-EAD responses by *Acleris comariana* to a natural pheromone gland extract. Bolded
640 response values with * were significant above the background response. Note the particularly significant
641 response (significant hit frequency 50.00 spikes/second) to the bioactive pheromone component (*E*)-
642 11,13-tetradecadienal (GC-FID peak 7).

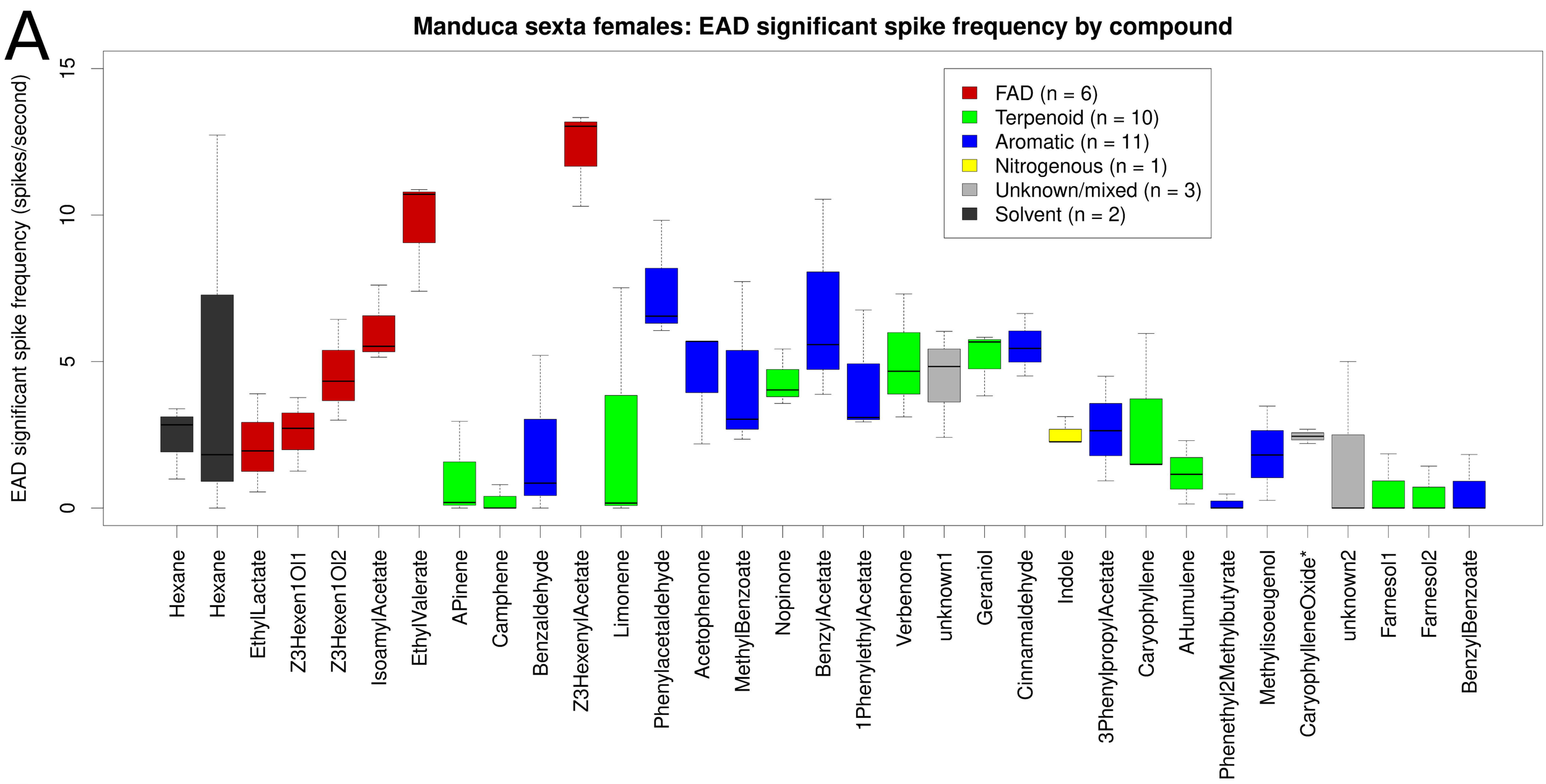
643 **Table S5:** Literature sources demonstrating previous electroantennogram activity or lack thereof in
644 *Manduca sexta* and *Bombus terrestris* to the compounds used in this study. Note the frequent
645 discrepancies in activity (especially for activity in *B. terrestris*), even for compounds with multiple
646 citations. Compounds with at least one active citation in either species are highlighted in bold.
647 Compounds shaded in yellow were EAD-active in more than 2/3 of *M. sexta* males or females and/or *B.*
648 *terrestris* worker individuals from this study.


GC-FID Signal
FID amplitude (pA)
0
400
800
Z3Hexen1Ol1
Z3Hexen1Ol2
IsoamylAcetate
EthylValerate
APinene
Camphene
Benzaldehyde
Z3HexenylAcetate
Limonene
Phenylacetaldehyde
4
5
6
7
8
Time (minutes)

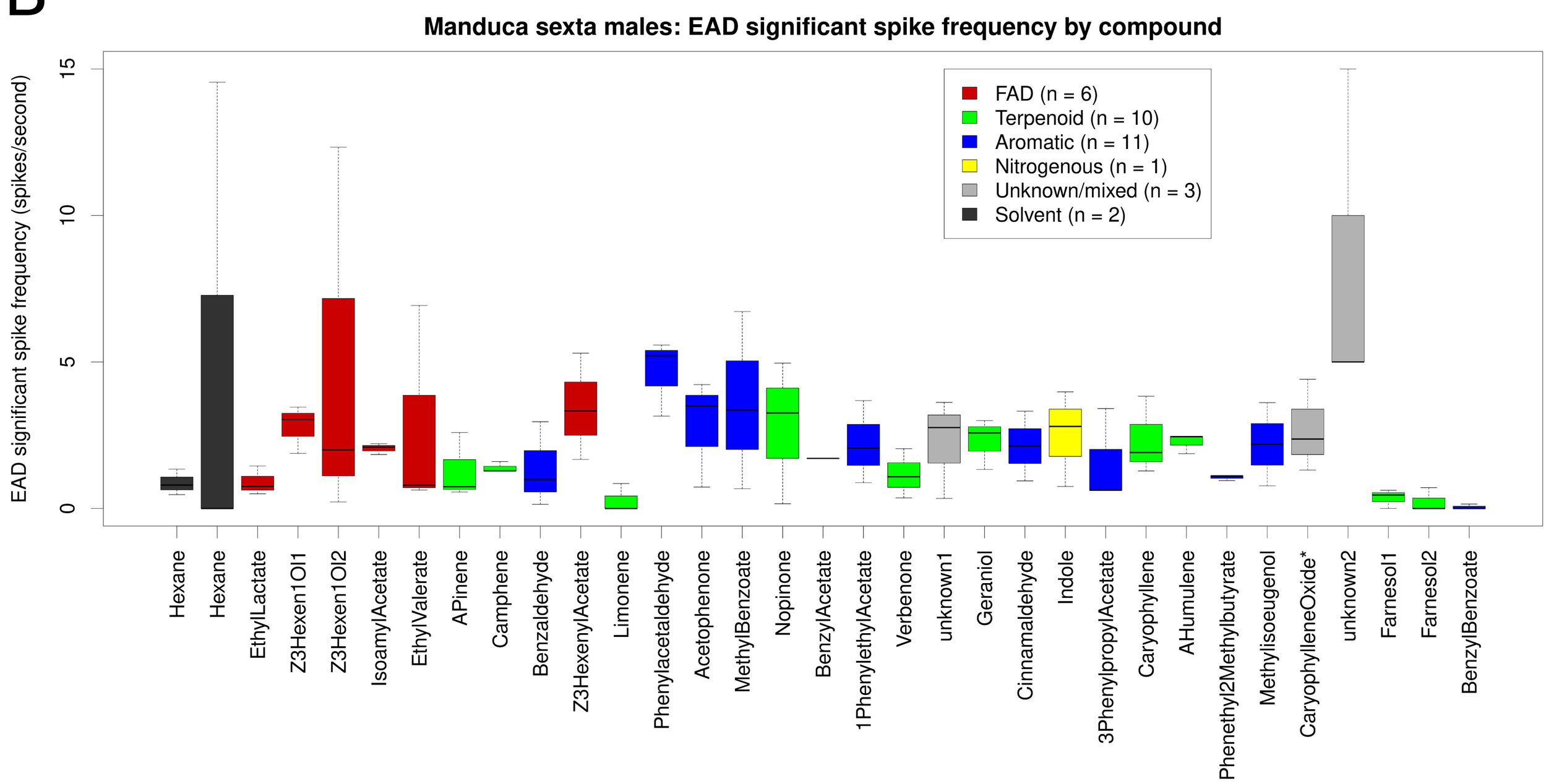

Moth 1 (female): raw EAD amplitude
Raw EAD amplitude (mV)
-1.5
-0.5
0.5
1.5
53
54
55
56
Time (minutes) (adjusted for GC-FID injection time)


Moth 1 (female): first forward difference of median-filtered antennal response
First forward difference
-0.2
0.0
0.1
0.2
53
54
55
56
Time (minutes) (adjusted for GC-FID injection time)

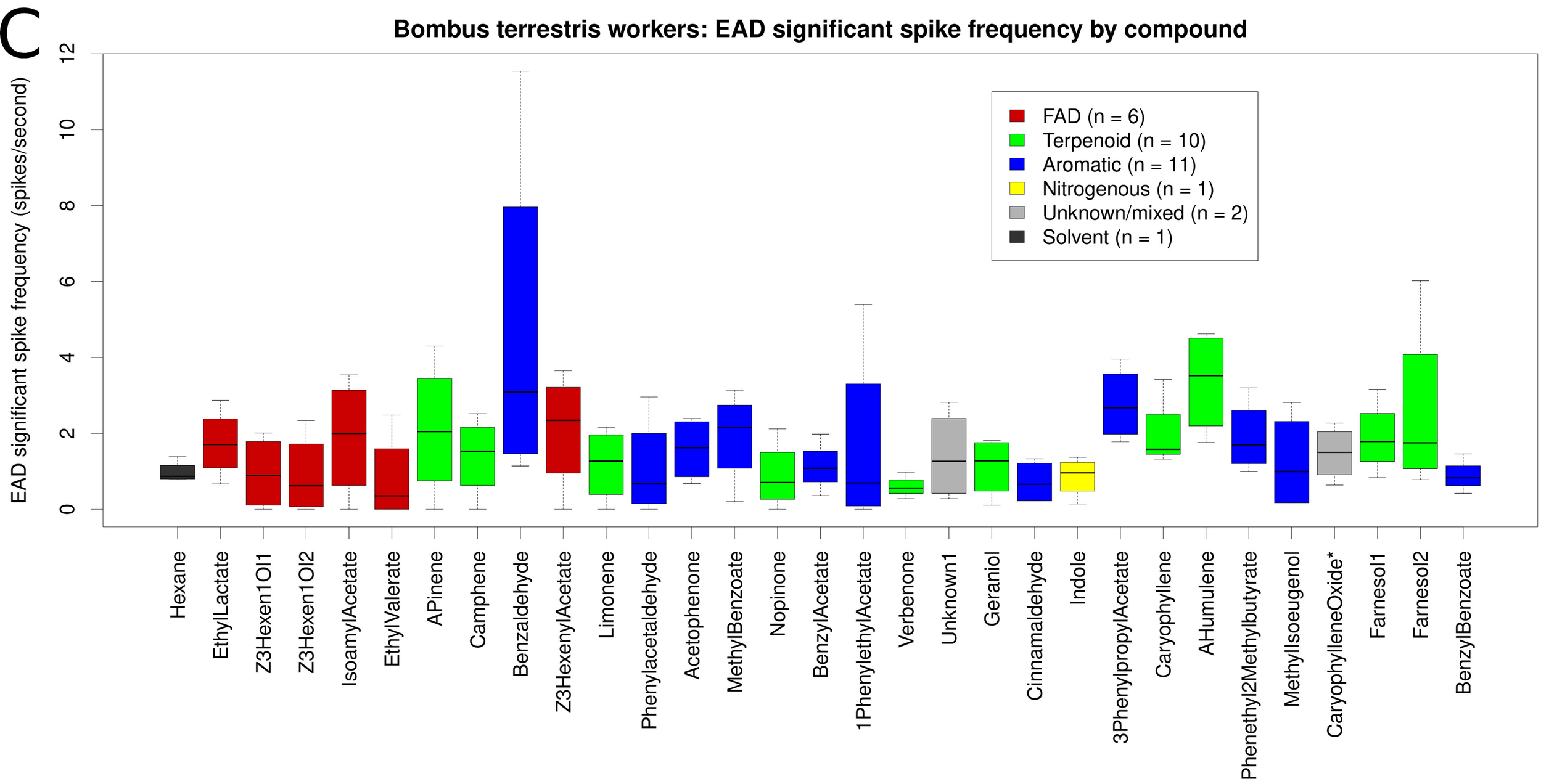
A

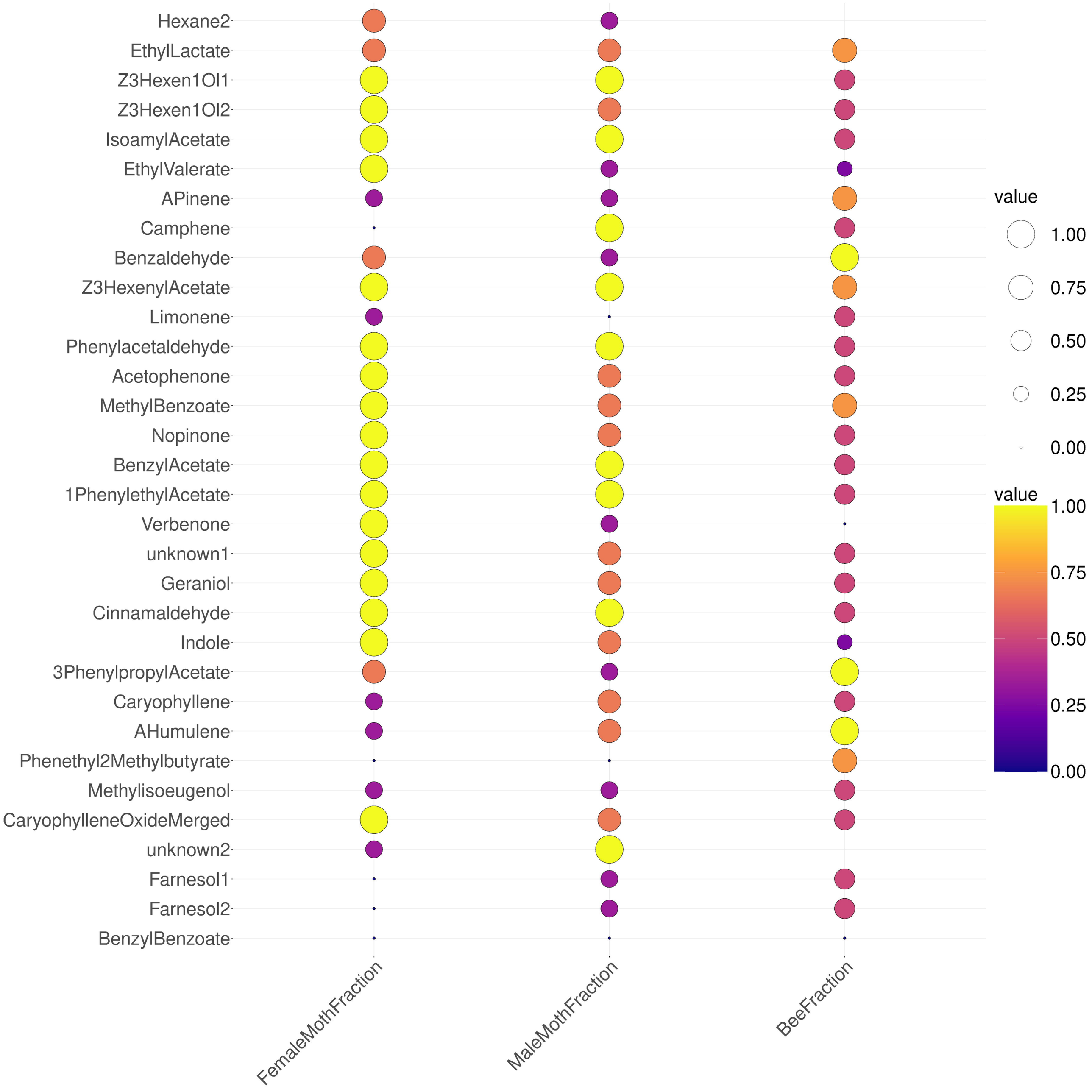

Manduca sexta females: EAD significant spike frequency by compound
EAD significant spike frequency (spikes/second)
0
5
10
15
FAD (n = 6)
Terpenoid (n = 10)
Aromatic (n = 11)
Nitrogenous (n = 1)
Unknown/mixed (n = 3)
Solvent (n = 2)
Hexane
Hexane
EthylLactate
Z3Hexen1Ol1
Z3Hexen1Ol2
IsoamylAcetate
EthylValerate
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

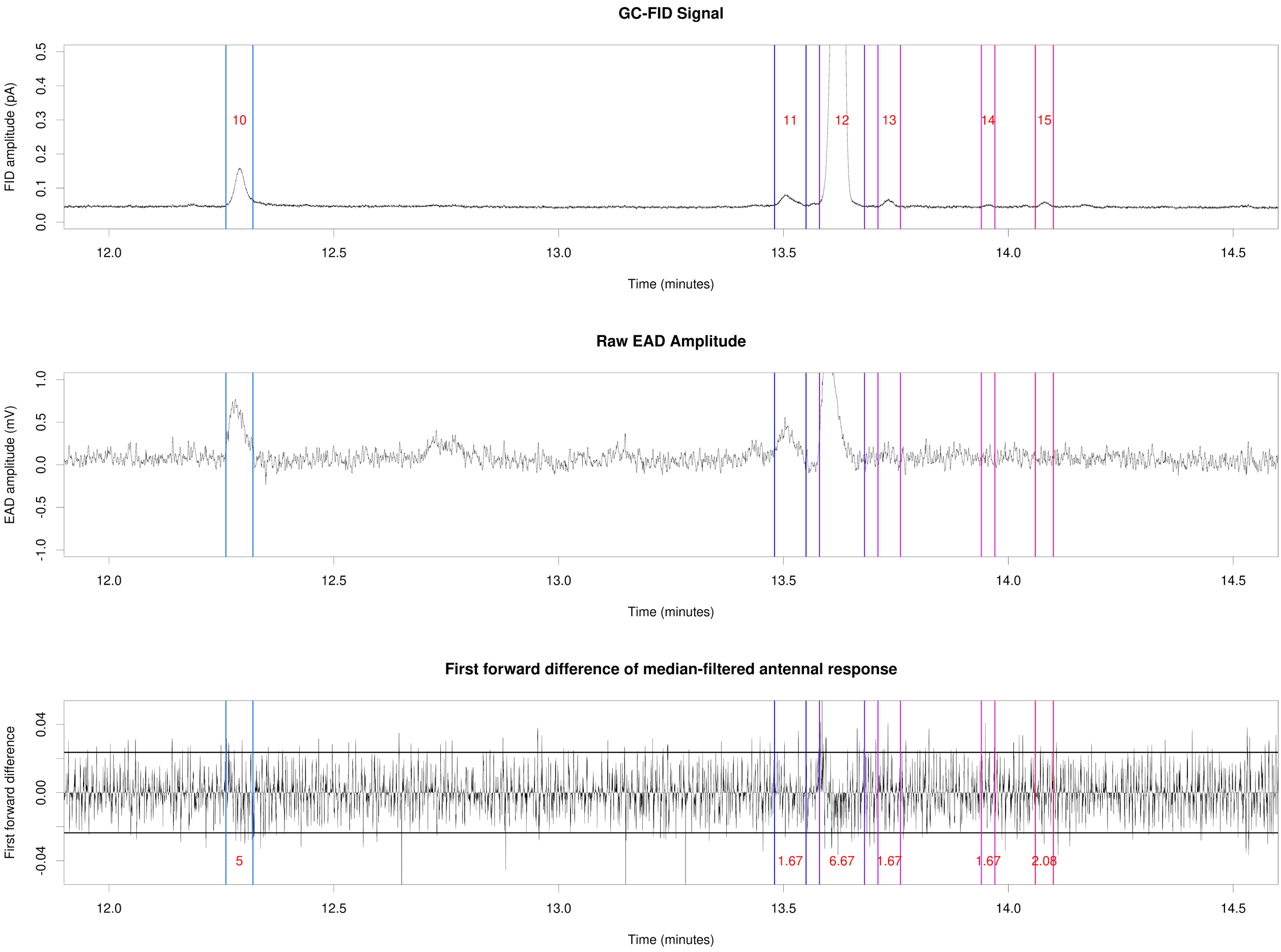
B

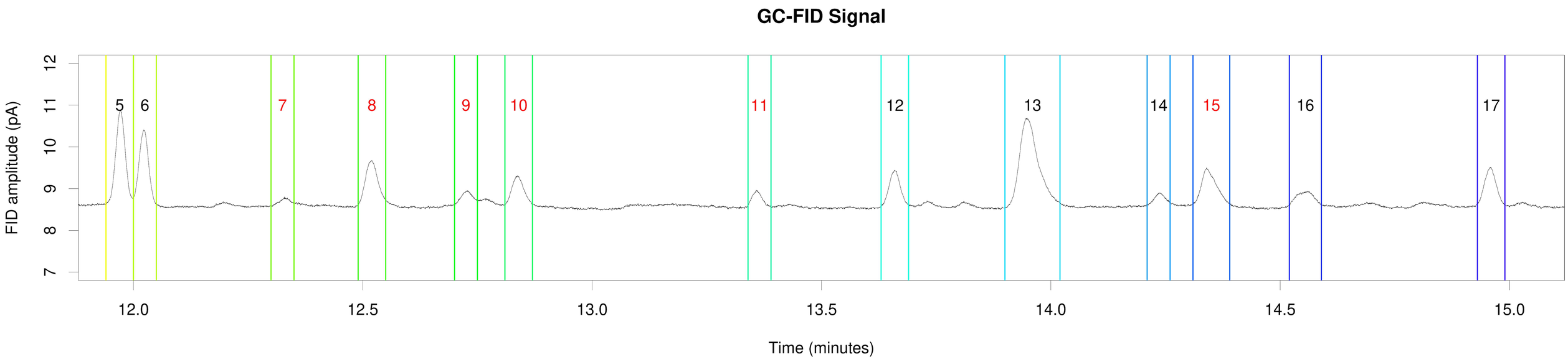

Manduca sexta males: EAD significant spike frequency by compound
EAD significant spike frequency (spikes/second)
0
5
10
15
FAD (n = 6)
Terpenoid (n = 10)
Aromatic (n = 11)
Nitrogenous (n = 1)
Unknown/mixed (n = 3)
Solvent (n = 2)
Hexane
Hexane
EthylLactate
Z3Hexen1Ol1
Z3Hexen1Ol2
IsoamylAcetate
EthylValerate
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

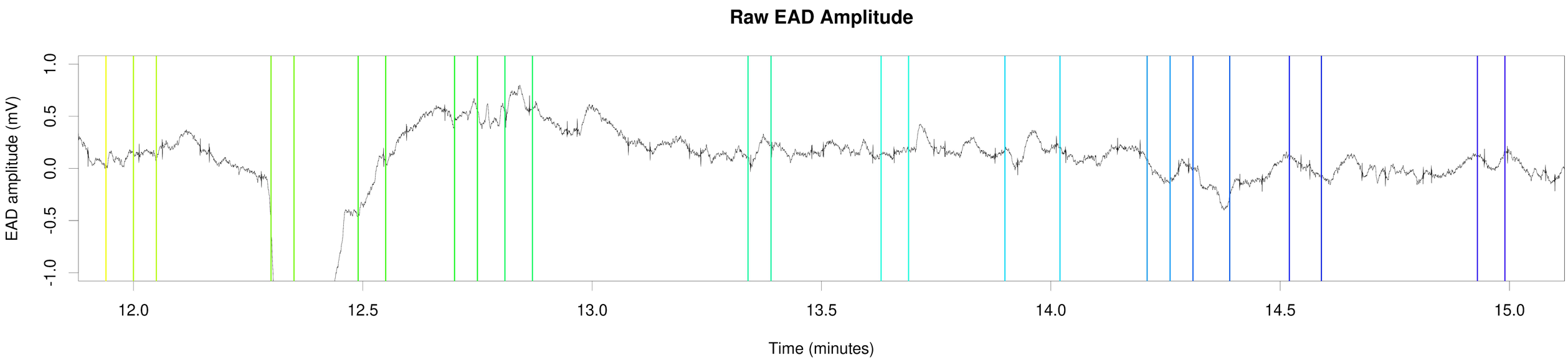
C

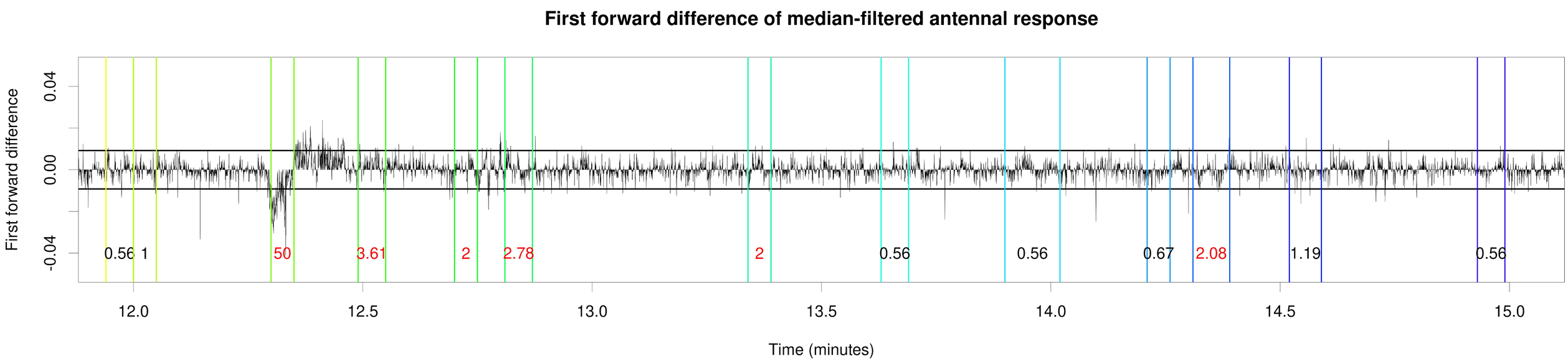

Bombus terrestris workers: EAD significant spike frequency by compound
EAD significant spike frequency (spikes/second)
0
2
4
6
8
10
12
FAD (n = 6)
Terpenoid (n = 10)
Aromatic (n = 11)
Nitrogenous (n = 1)
Unknown/mixed (n = 2)
Solvent (n = 1)
Hexane
EthylLactate
Z3Hexen1Ol1
Z3Hexen1Ol2
IsoamylAcetate
EthylValerate
APinene
Camphene
Benzaldehyde
Z3HexenylAcetate
Limonene
Phenylacetaldehyde
Acetophenone
MethylBenzoate
Nopinone
BenzylAcetate
1PhenylethylAcetate
Verbenone
Unknown1
Geraniol
Cinnamaldehyde
Indole
3PhenylpropylAcetate
Caryophyllene
AHumulene
Phenethyl2Methylbutyrate
MethylIsoeugenol
CaryophylleneOxide*
Farnesol1
Farnesol2
BenzylBenzoate


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
CaryophylleneOxideMerged
unknown2
Farnesol1
Farnesol2
BenzylBenzoate
FemaleMothFraction
MaleMothFraction
BeeFraction
value
1.00
0.75
0.50
0.25
0.00
value
1.00
0.75
0.50
0.25
0.00


GC-FID Signal
FID amplitude (pA)
0.0
0.1
0.2
0.3
0.4
0.5
10
11
12
13
14
15
12.0
12.5
13.0
13.5
14.0
14.5
Time (minutes)
Raw EAD Amplitude
EAD amplitude (mV)
-1.0
-0.5
0.0
0.5
1.0
12.0
12.5
13.0
13.5
14.0
14.5
Time (minutes)
First forward difference of median-filtered antennal response
First forward difference
-0.04
0.00
0.04
5
1.67
6.67
1.67
1.67
2.08
12.0
12.5
13.0
13.5
14.0
14.5
Time (minutes)


GC-FID Signal
FID amplitude (pA)
12
11
10
9
8
7
5
6
7
8
9
10
11
12
13
14
15
16
17
12.0
12.5
13.0
13.5
14.0
14.5
15.0
Time (minutes)


Raw EAD Amplitude
EAD amplitude (mV)
1.0
0.5
0.0
-0.5
-1.0
12.0
12.5
13.0
13.5
14.0
14.5
15.0
Time (minutes)


First forward difference of median-filtered antennal response
First forward difference
0.04
0.00
-0.04
0.56
1
50
3.61
2
2.78
2
0.56
0.56
0.67
2.08
1.19
0.56
12.0
12.5
13.0
13.5
14.0
14.5
15.0
Time (minutes)