



**HAL**  
open science

# Chemical Characterization of Young Virgin Queens and Mated Egg-Laying Queens in the Ant *Cataglyphis cursor*: Random Forest Classification Analysis for Multivariate Datasets

Thibaud Monnin, Florence Helft, Chloé Leroy, Patrizia D'ettore, Claudie Doums

## ► To cite this version:

Thibaud Monnin, Florence Helft, Chloé Leroy, Patrizia D'ettore, Claudie Doums. Chemical Characterization of Young Virgin Queens and Mated Egg-Laying Queens in the Ant *Cataglyphis cursor*: Random Forest Classification Analysis for Multivariate Datasets. *Journal of Chemical Ecology*, 2018, 44 (2), pp.127-136. 10.1007/s10886-018-0923-7. hal-02113130

**HAL Id: hal-02113130**

**<https://hal.science/hal-02113130>**

Submitted on 27 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 CHEMICAL CHARACTERIZATION OF YOUNG VIRGIN QUEENS AND MATED EGG-LAYING QUEENS IN  
2 THE ANT *Cataglyphis cursor*

3  
4 Thibaud MONNIN<sup>1\*</sup>, Florence HELFT<sup>1</sup>, Chloé LEROY<sup>4</sup>, Patrizia d'ETTORRE<sup>4</sup>, Claudie DOUMS<sup>2,3</sup>

5  
6 <sup>1</sup> *Institute of Ecology and Environmental Sciences of Paris UMR 7618, CNRS, Sorbonne Université,*  
7 *75005 Paris, France*

8 <sup>2</sup> *Institut de Systématique, Évolution, Biodiversité (ISYEB), UMR 7205, MNHN, CNRS, Sorbonne*  
9 *Université, EPHE, 75005 Paris, France*

10 <sup>3</sup> *EPHE, PSL Research University, 75005 Paris, France*

11 <sup>4</sup> *Laboratory of Experimental and Comparative Ethology, University of Paris 13, Sorbonne Paris Cité,*  
12 *Villetaneuse, France*

13  
14 \* *corresponding author. Thibaud.monnin@upmc.fr, +33 1 44 27 36 10*

15  
16  
17 **Abstract-** Social insects are well known for their extremely rich chemical communication, yet their  
18 sex pheromones remain poorly studied. We analysed the cuticular hydrocarbon profiles and Dufour's  
19 gland contents of queens of different ages and reproductive statuses (sexually immature gynes,  
20 sexually mature gynes, mated and egg-laying queens) and of workers in the thermophilic and  
21 thelytokous ant *Cataglyphis cursor*. Random forest classification analyses show that the four groups  
22 of individuals were well separated for both chemical sources, except mature gynes that clustered  
23 with queens for cuticular hydrocarbon and with immature gynes for Dufour's gland. Analyses carried  
24 out with two groups of females only allowed identifying candidate chemicals for queen signal and for  
25 sexual attractant. In particular, gynes produce more undecane in the Dufour's gland. This chemical is  
26 both the sex pheromone and the alarm pheromone of the ant *Formica lugubris*. It may therefore

27 possibly act as sex pheromone in *C. cursor*, and/or be involved in the restoration of monogyny that  
28 occurs rapidly following colony fission. Indeed, new colonies often start with several gynes and all  
29 but one are rapidly culled by workers, and this process likely involves chemical signals between gynes  
30 and workers. These findings open novel opportunities for experimental studies of inclusive mate  
31 choice and queen choice in *C. cursor*.

32

33 **Key Words-** Queen pheromone, sex pheromone, cuticular hydrocarbons, Dufour's gland, random  
34 forest.

35

36 **Acknowledgments-** None

37

38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63

## INTRODUCTION

Mate attraction and mate choice are essential features of living organisms (Clutton-Brock and Mcauliffe 2009, Jennions and Petrie 1997; Stearns and Hoekstra 2005;). They rely on different modalities of communication depending on species. For instance, birds predominantly rely on acoustic and visual signals (songs, plumage and courtship dances), mammals use acoustic, visual and chemical cues, and insects and arachnids mostly rely on tactile, acoustic and chemical signals. Social insects are well known for their extremely rich chemical communication. In ants, chemical trails often underline foraging, and cuticular hydrocarbon (CHC) profiles allow nestmate recognition (Hefetz 2007; van Zweden and d’Ettorre 2010) and the signalling of queen fertility (Holman et al. 2010; Holman et al. 2013; Le Conte and Hefetz 2008; Monnin 2006; van Oystaeyen et al. 2014). However, although ant semiochemicals are well studied in many regards their sex pheromones remain poorly studied (Ayasse et al. 2001).

Young, virgin, queens (hereafter gynes) attract males and advertise their sexual receptivity with volatile signals produced by the Dufour’s gland, the venom gland or the pygidial glands (reviewed in Ayasse et al. 2001; Hölldobler and Wilson 1990). To our knowledge, only two ant sex pheromones have been identified, in *Formica lugubris* (Walter et al. 1993) and in *Polyergus breviceps* (Greenberg et al. 2007). The chemical bouquet produced by a queen changes once she is mated, as reflected in modifications of the relative abundances of cuticular hydrocarbons within hours or a few days from mating in several ant species (Hora et al. 2008; Monnin 2006; Oppelt and Heinze 2009) and in the honeybee (Babis et al. 2014). This is because mated queens never re-mate (Boomsma 2013) hence need not attract males anymore whereas they need to advertise their fertility to colony members ( Foitzik et al. 2011; Hannonen et al. 2002; Holman et al. 2010; Holman et al. 2013; Keller and Nonacs 1993; Le Conte and Hefetz 2008; Monnin 2006; Ruel et al. 2013; van Oystaeyen et al. 2014).

64

65           We address this scarcity of knowledge of ant sex pheromones by examining the production  
66 of chemicals by queens of different ages and reproductive statuses in the thermophilic and  
67 thelytokous ant *Cataglyphis cursor*. Identifying queen signals would be particularly rewarding in this  
68 species where inclusive sexual selection and culling of supernumerary queens occur and likely involve  
69 chemical signals. In *C. cursor*, gynes do not fly. They mate at the nest entrance, where foreign males  
70 gather (Cronin et al. 2011; Lenoir et al. 1988). Gynes show no mate choice, but workers harass males  
71 seeking to mate (Helft et al. 2015; Helft et al. 2016). How males are attracted to the nest entrance  
72 and how they identify sexually receptive gynes are unknown in *C. cursor*, but they likely rely on  
73 chemical signals. Newly mated gynes then disperse with workers to found new colonies (dependent  
74 colony foundation, Chéron et al. 2011a; Cronin et al. 2013; Lenoir et al. 1988). Many new colonies  
75 are temporarily polygynous, but this state is transitory and monogyny is restored within a few days  
76 (Chéron et al. 2011a; Cronin et al. 2012). How monogyny is restored is unknown, but it likely entails  
77 some degree of competitive chemical signalling of newly-mated queens vying for colony ownership.  
78 Identifying potential queen signals would therefore allow designing precise key bioassays and studies  
79 on inclusive sexual selection and queen competition in this interesting ant species.

80

81           Queen chemistry has not been studied in *C. cursor* and very little is known from other  
82 *Cataglyphis* species (Dahbi and Lenoir 1998; van Oystaeyen et al. 2014), even though worker  
83 chemistry has been studied in *C. cursor* ( Hefetz and Lenoir 1992; Nowbahari et al. 1990) and in other  
84 *Cataglyphis* species (Ali et al. 1988; Dahbi et al. 2008; Gökçen et al. 2002; Hefetz and Lenoir 1992).  
85 The CHC profile of three mature, egg-laying, *C. iberica* queens was compared with that of their  
86 workers. It revealed no qualitative differences but higher relative abundances of three linear alkanes  
87 ( $n\text{-C}_{27}$ ,  $n\text{-C}_{28}$ ,  $n\text{-C}_{29}$ ) and one dimethylalkane (dime $\text{C}_{29}$ ) and a lower relative abundance of two  
88 methylalkanes (11- and 13-me $\text{C}_{27}$ ) in queens (Dahbi and Lenoir 1998). The treatment of orphaned  
89 workers with a mixture of linear alkanes and methylalkanes ( $n\text{-C}_{27}$ ,  $n\text{-C}_{29}$  and 3-me $\text{C}_{29}$ ) repressed their

90 ovarian development, suggesting that these chemicals function as queen pheromones (van  
91 Oystaeyen et al. 2014).

92

93 To our knowledge, no gyne (i.e. young virgin queen) has been studied in *Cataglyphis*. We  
94 analyse here the cuticular hydrocarbon (CHC) profiles and Dufour's gland (DG) contents of egg-laying  
95 queens, sexually receptive gynes, gynes too young to be sexually receptive and workers of *C. cursor*.  
96 Our hypothesis is that chemicals over-expressed in egg-laying queens and in sexually receptive gynes  
97 may be candidate components of queen pheromones and sexual pheromones, respectively.

98

99

## 100 METHODS AND MATERIALS

101

102 Sixteen complete colonies of *C. cursor* were collected in Argelès-sur-Mer, Southern France  
103 (42.5722508°N, 3.043670°E), on the beginning of March 2014 (1 colony) and March 2015 (15  
104 colonies) (Table 1). This is right after the end of hibernation and immediately prior the onset of  
105 sexual production in April-May, with mating taking place in May-June (Chéron et al. 2011a; Cronin et  
106 al. 2011). Colonies contained  $642 \pm 496$  workers (mean  $\pm$  SD, range 168-1857) and were set up in the  
107 laboratory within a week of the collection. Each colony was housed in a plaster nest contained in a 35  
108 x 25 x 12 cm box as foraging arena, kept under a 12:12 day:night light cycle and 25-28°C, provided  
109 with sugar and water *ad libitum*, and fed mealworm and crickets twice to thrice a week. Three  
110 colonies were queenless by late April, when we started monitoring sexual production (Table 1). We  
111 retained them because *C. cursor* workers produce new queens by thelytokous parthenogenesis when  
112 orphaned (Cagniant 1979; Chéron et al. 2011b).

113

114 Only four colonies, including the three orphaned colonies, produced sexuals in the laboratory  
115 (Table 1). These four colonies produced gynes, but only two colonies (one queenright and one

116 orphaned) produced males. Note that orphaned colonies mainly produce diploid males who behave  
117 like haploid males and can mate and fertilise the queen (Doums et al. 2013b). Males were removed  
118 upon emergence to prevent them from dispersing to other colonies and mate, hence gynes remained  
119 virgin. We sampled 25 gynes between 11<sup>th</sup> of May and 5<sup>th</sup> of June 2015. Twelve were sampled when  
120 they were one to five days old (mean age of three days, hereafter gynes G0) while the remaining 13  
121 were sampled at eight to 11 days old (mean age of 10 days, gynes G1). Gynes G0 and G1 are sexually  
122 immature and sexually mature, respectively, and this was confirmed by the shedding of wings and  
123 dissections (see results). They originated from four and three sexuals-producing colonies,  
124 respectively (Table 1). This introduces some pseudo replication in our dataset. This is common in  
125 chemical studies of social insect sexuals because of the difficulties for obtaining samples from many  
126 independent colonies. Our study focuses on chemicals involved in behaviours that are not expected  
127 to differ in any consistent way between colonies (mate attraction, competition between young  
128 queens for colony “ownership”) hence these chemicals should not be affected by colony effects and  
129 our result should be robust. Gynes were immediately stored at -80°C when collected, until later  
130 dissection and chemical extraction. In addition, 13 queens (Q) and 16 workers (W) were collected on  
131 the 9<sup>th</sup> of June hence totalling 54 individuals (Table 1).

132

133           Chemical extraction and dissections were carried out at the end of June. Each ant was  
134 individually retrieved from the freezer immediately before its chemical extraction and dissection. We  
135 recorded whether gynes were winged or dealated as a cue of sexual maturity independent from our  
136 age-based groups of Gynes G0 and G1. Wing loss is indicative of mating in the field where 83.5% of  
137 virgin gynes vs. 16% of mated gynes were winged (Doums et al. 2013a). We then separated the  
138 gaster from the rest of the body. The gasterless body was immersed in 200 µl of pentane for 5 min in  
139 order to extract the cuticular hydrocarbons. At the same time, we dissected the gaster to collect the  
140 Dufour’s gland, which was extracted in 100 µl of pentane. Lastly, we dissected the reproductive  
141 apparatus and recorded mating status and ovarian development.

142

143 *Chemical Analysis*

144

145 Samples were analysed by gas chromatography-mass spectrometry (GC-MS). Each sample was  
146 evaporated at ambient temperature under an extractor hood and re-diluted with 70µl of pentane.  
147 Two µl of the solution were injected using an Agilent G4513A Automatic Liquid Sampler into an  
148 Agilent Technologies 7890A gas chromatograph (capillary column: Agilent HP-5MS, 30 m × 0.25 mm ×  
149 0.25 µm; split–splitless injector; carrying helium gas at 1 mL/min) coupled with an Agilent 5975 C  
150 mass spectrometer with 70 eV electron impact ionization. The injector was set at 280°C and the  
151 temperature programme was as follows: an initial hold at 70°C for 1 min, then 70°C to 175°C at  
152 10°C.min<sup>-1</sup>, and up to 320°C at 5°C.min<sup>-1</sup> and hold for 3 min.

153

154 Peak areas were integrated with MSD ChemStation software version E.02.01.1177 (Agilent  
155 technologies). We used the best profile for each chemical source (CHC, DG) and each group of  
156 individuals (W, G0, G1, Q) to select and identify the peaks to include in our analyses. We identified  
157 chemical compounds by their retention time, fragmentation pattern and comparisons with published  
158 results on *Cataglyphis* species including *C. cursor* (Ali et al. 1988; Dahbi et al. 2008; Gökçen et al.  
159 2002; Hefetz and Lenoir 1992; Keegans et al. 1992; Nowbahari et al. 1990).

160

161 *Statistical Analysis*

162

163 The results were analysed separately for each chemical source (CHC, DG). We used random forest  
164 classification analysis (RF) (Liaw and Wiener 2002) to see whether the four groups (egg-laying  
165 queens, sexually mature gynes, sexually immature gynes, workers) could be discriminated and to  
166 identify peaks that best discriminate groups. RF is a method increasingly used in chemical studies,  
167 including the chemical ecology of social insects (Jansen et al. 2016; Mitra et al. 2014; Oliveira et al.



168 2016). It is a tree-based machine learning tool for analysing multivariate dataset with more variables  
169 than individuals (Goldstein et al. 2011), and this is typically the case in chemical analyses. In addition,  
170 RF deals well with multicollinearity, assumes no specific distribution of the variables and does not  
171 require transformation of constrained data (such as chemical compounds) (Ranganathan and Borges  
172 2010).

173

174 We used R v3.2.2 (R Core Team 2016) with Rstudio v1.0.44 (RStudio Team 2016). We used  
175 the relative abundances of peaks (i.e. proportions each peak represents relative to the sum of all  
176 peaks) computed from the raw dataset, including all the zero values. RF were ran using the package  
177 randomForest (Liaw and Wiener 2002). We performed RF using 10,000 trees to determine the  
178 importance of each peak in classifying the individuals in the four groups. We used the default value  
179 for the number of input variables randomly selected to build each node of the tree (denoted by  
180 *mtry*), that is the squared root of the number of variables (seven in our case). We also used the  
181 default value (33%) for defining the set of observations not used for building the current tree, i.e. the  
182 out of bag sample (OOB sample). This OOB sample is used to generate the confusion matrix and OOB  
183 error rate (i.e. rate of unsuccessful assignment of the OOB samples to the correct group). A low OOB  
184 error rate indicates a high ability of the indicator variables to predict the group of origin of the  
185 individual. The separation of the groups was visualized using a multidimensional scaling plot based  
186 on a proximity matrix of the individuals on the RF (i.e. the number of times they end up on the same  
187 leaf in the trees) (function MDSplot).

188

189 For each chemical source, as a second step, we tried to identify the peaks that are the most  
190 important for queen recognition by comparing queens vs. workers, and for sexual attraction by  
191 comparing gynes G1 vs. gynes G0 and gynes G1 vs. queens. To do so, we ran three additional RF  
192 analyses, each contrasting only the two groups of interest, and used the package VSURF (Genuer et  
193 al. 2015) to identify the peaks important in discriminating the two groups. The measure of variable

194 importance is based on the decrease of classification accuracy when the observed values of the given  
195 variable are randomly permuted in the OOB samples (Genuer et al. 2015). The selection procedure is  
196 based on three steps. The first (thresholding step) ranks variables according to a score of importance,  
197 and eliminates variables with a score lower than a threshold value. The second selects the variables  
198 needed to best discriminate the groups of interest. Embedded RF models are grown, starting with a  
199 RF build with only the most important variable and finishing with a RF with all the variables retained  
200 in the previous step, and the best model is selected on the basis of the means OOB error. The third  
201 step refines the previous selection for a predictive purpose and was therefore not used here. As  
202 above, the RF models were run with 10,000 trees and a sample of seven peaks for each node. The  
203 selection procedure for each comparison was run 20 times to assess the stability of the variable  
204 selection process. For each comparison, the chemical compounds were therefore ranked by their  
205 level of importance in discriminating the two groups and we averaged their ranks over the 20 runs of  
206 the variable selection process.

207

208

209

## RESULTS

210

211 The dissections confirmed that all gynes were virgin, as expected since males were removed from  
212 colonies to prevent mating. Gynes G0 were all winged and had inactive ovaries without oocytes or  
213 with a few small and non-yolky oocytes at most. Seven gynes G1 were dealated (three partially and  
214 four fully) and six were winged. Gynes G1 had slightly more developed ovaries than gynes G0: five  
215 had small non-yolky oocytes while the remaining eight had small but yolky oocytes. This supports our  
216 hypothesis that gynes G1 were sexually mature because dealation is a good indicator that a gyne is  
217 mated in nature (Doums et al. 2013a). All queens were mated and had active ovaries with yellow  
218 bodies, except the queen of colony 1268 which showed regressing ovaries indicative of aging.

219

220 Table S1 lists the compounds identified from the cuticle (60 peaks) and Figure S1 provides  
221 representative chromatograms for each group of individuals. Similarly, Table S2 lists the compounds  
222 from the Dufour's gland (36 peaks) and Figure S2 shows representative chromatograms. The  
223 compounds identified include alkanes, methylalkanes and alkenes. In addition, some acetates and  
224 alcohols were also found in the DG.

225

### 226 *Cuticular Hydrocarbons*

227

228 The RF analysis with the four groups of individuals shows that workers were clearly separated from  
229 sexual females (gynes and queens) along the first axis, while the three groups of sexual females were  
230 separated along the second axis (Figure 1). Gynes G0 were clearly separated from queens and most  
231 gynes G1 clustered with queens, although two clustered with gynes G0. That is, there was a gradual  
232 transition of the CHC profile from G0 to G1 and queens. The OOB error rate (i.e. individuals not  
233 correctly assigned to their group) was 12.96% and most misclassifications occurred between gynes  
234 G1 and either gynes G0 (2 errors) or queens (4 errors, Table S3).

235

236 In order to identify the compounds most important for differentiating two given types of  
237 females (Q vs. W; G1 vs. G0; G1 vs. Q), we ran three additional RF analyses considering only two  
238 groups of individuals at a time, for both CHC and DG. We selected the variables (compounds) of  
239 importance for the discrimination using the package VSURF and ran each variable selection process  
240 20 times to assess the stability of the selection process (see Methods and Materials). The results  
241 show that the importance of chemical compounds was stable over the 20 runs. The first step of the  
242 variable selection procedure (thresholding step) retained 45 to 51 peaks out of 60 peaks (i.e. 75-  
243 85%). Most compounds were consistently ranked with the same importance, and when variation  
244 occurred it was small. The variable selection procedure was reliable for the three paired

245 comparisons, and thus for the sake of clarity we present the 20 peaks with the highest averaged  
246 ranks over the 20 runs (Figure 2).

247

248 The Q/W comparison aimed at identifying chemicals that may be queen signals. The OOB  
249 error rate was much lower (3.45%) than when considering the four groups. The 20 peaks with the  
250 highest averaged rank of importance show a clear queen/worker difference in terms of the weight of  
251 their CHCs, with the heavier compounds being more abundant in queens than in workers (Figure 2A).  
252 The second step of the variable selection procedure retained a single compound as needed for  
253 obtaining the best discrimination (Peak 26 = 3-meC<sub>27</sub>). However, this chemical was more abundant in  
254 workers than in queens (Table S1) hence is unlikely to be a queen signal.

255

256 The comparison gynes G1 vs. gynes G0 aimed at identifying compounds differentiating  
257 sexually mature and sexually immature gynes. In agreement with the results obtained when  
258 considering the four groups of individuals (Figure 1), gynes G1 and G0 were well differentiated (OOB  
259 error rate of 8%). The second step of the selection procedure identified three chemicals as necessary  
260 for best discrimination (P35 = *n*-C<sub>29</sub>, P44 = 14-meC<sub>30</sub> + 8,12-dimeC<sub>30</sub> and P50 = 11,13-dimeC<sub>31</sub>), among  
261 which one was more abundant in gynes G1 and two were more abundant in gynes G0 (Figure 2B).

262

263 Last, we compared gynes G1 and queens. Again, in agreement with the results when  
264 considering the four groups of individuals (Figure 1), they were rather poorly differentiated (OOB  
265 error rate of 15%). Nine of the 20 peaks with the highest averaged rank of importance were more  
266 abundant in queens than in gynes G1, and these comprised the four heaviest and the five lightest  
267 compounds (Figure 2C). Three peaks were required for the best discrimination (P53 = *n*-C<sub>33</sub>, P47 = *n*-  
268 C<sub>31</sub> and P54 = 15+17-meC<sub>33</sub>), and they were more abundant in queens and among the four heaviest  
269 compounds.

270

271 *Dufour's Gland*

272

273 The RF analysis with the four groups of individuals shows that the first axis separates workers from  
274 sexual females, and that the second axis separates sexual females (Figure 1). Gynes G0 and G1  
275 clustered together and were well separated from queens. That is, there was a gradual transition of  
276 the DG content from G0 and G1 to queens, but while gynes G1 clustered with queens for CHC they  
277 clustered with gynes G0 for DG. The OOB error rate was 11.11%, and most misclassification occurred  
278 between gynes G0 and G1 (DG: 5 errors, Table S3).

279

280 We performed the same variable selection procedure than for CHC, using RF to identify the  
281 compounds most important in differentiating two given types of females. As for CHC, the variable  
282 selection procedure was stable over the 20 runs. The first step of the procedure (thresholding step)  
283 retained 24 to 35 peaks out of 36 peaks (67-97%) and for the sake of clarity we present the 20 peaks  
284 with the highest averaged ranks over the 20 runs (Figure 3).

285

286 The RF comparing queens and workers yielded a low OOB error rate of 3.45%, in agreement  
287 with the results obtained when including the four groups. As for CHC, there was a very clear  
288 queen/worker difference in terms of the weight of chemicals, with all but one of the heavier  
289 compounds being more abundant in queens than in workers (Figure 3A). The second step retained  
290 one compound only as needed for best discrimination, but this compound varied across runs (it was  
291 either peak 27 = C<sub>27:1</sub>, P31 = C<sub>29:1</sub> or P33 = C<sub>31:2</sub>). This is because these three compounds were  
292 correlated hence the information they provide for discriminating the two groups was redundant. It is  
293 worth noticing that the three compounds are structurally related (alkene) and more abundant in  
294 queens than in workers (Table S2).

295

296 The comparison of gynes G1 and G0 yielded a high OOB error rate (20%), which is also in  
297 agreement with the results obtained when considering the four groups of individuals where G1 and  
298 G0 clustered together (Figure 1). The second step of the selection procedure showed that three  
299 peaks are needed for the best discrimination (P15 = C<sub>19:2</sub>, P12 = me-C<sub>17</sub> and P21 = C<sub>23:1</sub>), including two  
300 more abundant in gynes G1 and one more abundant in gynes G0 (Figure 3B).

301

302 Last, we compared gynes G1 and queens. They were perfectly differentiated (OOB error rate  
303 of 0%), in agreement with the comparison using the four groups of individuals (Figure 1). Most of the  
304 peaks more abundant in queens were heavy compounds (Figure 3C). The second step selected three  
305 peaks (P8 = tridecanol, P7 = dodecanyl acetate and P1 = *n*-C<sub>11</sub>). All were more abundant in gynes G1  
306 and among the light compounds (Figure 3C), which is consistent with their putative use as volatile  
307 attractants of foreign males for mating.

308

309

310

311

## DISCUSSION

312

313 We examined the changes in chemical production that *C. cursor* sexual females undergo as they age  
314 and mate, by comparing virgin and sexually immature queens (gynes G0), virgin and sexually mature  
315 queens (gynes G1) and mated and egg-laying queens. We included workers as a control and analysed  
316 the chemical compounds from the cuticle and the Dufour's gland. As expected, the four groups of  
317 individuals were overall well discriminated, for the two chemical sources. Subsequent analyses  
318 contrasting two types of females each (queens vs. workers, gynes G1 vs. gynes G0 and gynes G1 vs.  
319 queens) pinpointed the chemical compounds that are good candidate as chemical signals.

320

321

Compounds that were of high importance for discriminating queens from workers and that were  
more abundant in the former may be queen signals. Similarly, compounds that were important in

322 discriminating gynes G1 from G0 or from queens and that were more abundant in gynes G1 may be  
323 male attractants or chemicals involved in the restoration of monogyny.

324

325         The branched alkane 3-meC<sub>27</sub> was important in discriminating queen and worker CHC  
326 profiles. This is in apparent agreement with previous studies that have found 3-methylalkanes to be  
327 characteristic of queen vs. worker CHC profiles in several social insect species (3-meC<sub>29</sub> in *C. iberica*:  
328 Oystaeyen et al. 2014; from 3-meC<sub>25</sub> to 3-meC<sub>35</sub> in several *Lasius* species: Holman et al. 2013; Oi et al.  
329 2015). However, 3-meC<sub>27</sub> was less abundant in *C. cursor* queens than workers, whereas in contrast 3-  
330 methylalkanes are more abundant in *C. iberica* and *Lasius* queens than workers (Holman et al. 2013;  
331 van Oystaeyen et al. 2014). It is therefore doubtful that 3-meC<sub>27</sub> could be (a component of) the  
332 queen signal in *C. cursor*. Other compounds that were both important in discriminating queen and  
333 worker CHC profiles and more abundant in the former may constitute the queen signal. These  
334 include linear alkanes (*n*-C<sub>31</sub> and *n*-C<sub>33</sub>), methylalkanes (13+15-meC<sub>31</sub> and 15+17-meC<sub>33</sub>) and alkenes  
335 (C<sub>23:1</sub>, C<sub>31:1</sub> and C<sub>33:1</sub>) (Figure 2 and Table S1). In addition, *C. cursor* queens also differed from workers  
336 by producing more alkenes (C<sub>27:1</sub>, C<sub>29:1</sub> and C<sub>31:2</sub>) in the Dufour's gland. Chemicals of these types have  
337 been identified as putative queen signals in several other social insect species (e.g. Le Conte and  
338 Hefetz 2008; Oi et al. 2016; van Zweden et al. 2014) including the congeneric *C. iberica* (Dahbi and  
339 Lenoir 1998; van Oystaeyen et al. 2014). It is now relatively consensual that queen signals are honest  
340 (Heinze and d'Etorre 2009; Keller and Nonacs 1993), but whether they are largely conserved across  
341 social insect species ( Oi et al. 2016; van Oystaeyen et al. 2014) or variable (Amsalem et al. 2015;  
342 Evison et al. 2012; Smith et al. 2016) remains debated. The activity of the queen-characteristic  
343 compounds presently identified in *C. cursor* should be determined with an adequate bioassay, such  
344 as whether they inhibit (delay or decrease) egg-laying by orphaned workers (e.g. Oi et al. 2016). The  
345 queen is multiply mated in *C. cursor* and it would also be worth investigating the sensitivity of worker  
346 patriline to queen pheromones, to test for instance whether it correlates negatively with their

347 propensity to reproduce in presence of their queen ( i.e. behave as cheaters, Doums et al. 2013b) or  
348 to enter foreign colonies and reproduce ( i.e. behave as social parasite, Chéron et al. 2011b).

349

350 CHC profiles also discriminated well gynes G0 from gynes G1, and this likely stems from the  
351 fact that gynes G0 had a callow-like chemical profile while gynes G1 had already started producing a  
352 queen profile given that they had started developing their ovaries. In contrast, DG contents did not  
353 differentiate between the two categories of gynes, which suggests that CHC and DG contents mature  
354 at different speeds. Complementary studies are necessary to get a better understanding of the  
355 maturation of gynes, for instance by using additional age groups of gynes to cover a larger range of  
356 age differences, or by sampling gynes performing sexual calling in the field (as in Walter et al. 1993).

357

358 Although the comparison between gynes G1 and gynes G0 did not allow identifying putative  
359 sex pheromones, the comparison of gynes G1 with queens was more informative. CHC profiles did  
360 not discriminate well these two types of females. This may be because gynes G1 had started  
361 developing their ovaries and producing queen pheromones. However, the Dufour's gland content  
362 allowed a perfect discrimination of gynes G1 and queens, in agreement with our expectation that  
363 virgin and sexually mature gynes G1 produce male attractants whereas mated and egg-laying queens  
364 do not. Undecane ( $n\text{-C}_{11}$ ) was one of three components allowing this discrimination and it was more  
365 abundant in gynes G1 than in queens. Undecane is the main component of the sex pheromone of the  
366 ant *Formica lugubris* (Walter et al. 1993), opening up the possibility that it may be (a component of)  
367 the sex pheromone of *C. cursor*. The effect of undecane could be experimentally tested by recording  
368 males' responses to synthetic chemical deposited on filter paper, dummies (e.g. glass beads or dead  
369 individuals) or live individuals (workers or non-receptive queens). *C. cursor* workers are involved in  
370 inclusive sexual selection (Helft et al. 2015; Helft et al. 2016) and it would be interesting to test their  
371 possible contribution to male attraction. Workers use DG secretions for marking the surroundings of  
372 the nest entrance (Mayade et al. 1993) and by doing so they may release undecane and contribute to



373 attract males. Indeed, although each worker harbours little undecane in her DG (Table S2), workers  
374 may collectively deposit significant amounts.

375

376 Another possibility is that undecane is involved in the restoration of monogyny during the  
377 process of colony fission. Indeed, undecane has two functions in *F. lugubris*: sex pheromone in  
378 queens and alarm pheromone in workers (Walter et al. 1993). In *C. cursor* new colonies are founded  
379 by colony fission. They often start with several gynes but monogyny is restored within a few days  
380 (Chéron et al. 2011a) by fights between young queens and by attacks by workers (pers. obs.).  
381 Undecane could be a signal allowing physiologically more fertile gynes to mark less fertile gynes and  
382 thereby target worker attacks against the latter, as occurs in some other species in comparable  
383 contexts (de Souza et al. 2017; Monnin et al. 2002; Smith et al. 2012). The involvement of undecane  
384 in intra-colonial aggressions could be experimentally tested by recording workers' response to the  
385 deposit of synthetic chemical on live individuals (workers or non-receptive queens). A preliminary  
386 study indeed suggests that undecane triggers worker aggressions against male. When a worker and a  
387 male from another colony were placed together in a small box containing a filter paper, the worker  
388 attacked the male in 6 out of 18 replicates when the filter paper contained undecane whereas she  
389 did not when the filter paper was blank (0 out of 18 trials, Fisher exact test:  $p = 0.019$ ).

390

391 Our results show that the complex chemical bouquet of *C. cursor* queens changes according  
392 to age, mating status and reproductive activity, as occurs in other species where mating triggers  
393 chemical switches (Babis et al. 2014; Hora et al. 2008; Oppelt and Heinze 2009). These changes are  
394 related to mate attraction and to queen signalling towards workers, and the identification of  
395 candidate chemicals is a first step towards identifying pheromones. These would in turn allow  
396 exciting studies of the chemical determinants of mating site selection by males, restoration of  
397 monogyny (i.e. queen choice) by workers, and the link between sensitivity to queen pheromone and  
398 cheating and social parasitism.

399

400

401

## REFERENCES

- 402 Ali MF, Billen JPJ, Jackson BD, Morgan ED (1988) Secretion of the dufour glands of two african desert  
403 ants, *Camponotus aegyptiacus* and *Cataglyphis savignyi* (Hymenoptera: Formicidae). *Biochem*  
404 *Syst Ecol* 16:647–654. doi: 10.1016/0305-1978(88)90079-8
- 405 Amsalem E, Orlova M, Grozinger CM (2015) A conserved class of queen pheromones? Re-evaluating  
406 the evidence in bumblebees (*Bombus impatiens*). *Proc R Soc B* 282:20151800. doi:  
407 10.1098/rspb.2015.1800
- 408 Ayasse M, Paxton RJ, Tengö J (2001) Mating behavior and chemical communication in the order  
409 Hymenoptera. *Annu Rev Entomol* 46:31–78.
- 410 Babis M, Holman L, Fenske R, et al (2014) Cuticular lipids correlate with age and insemination status  
411 in queen honeybees. *Insectes Soc* 61:337–345. doi: 10.1007/s00040-014-0358-2
- 412 Boomsma JJ (2013) Beyond promiscuity: mate-choice commitments in social breeding. *Philos Trans R*  
413 *Soc Lond B* 368:20120050. doi: 10.1098/rstb.2012.0050
- 414 Cagniant H (1979) La parthénogenèse thélytoque et arrhénotoque chez la fourmi *Cataglyphis cursor*  
415 Fonsc. (Hym. Form.). Cycle biologique en élevage des colonies avec reine et des colonies sans  
416 reine. *Insectes Soc* 26:51–60. doi: 10.1007/BF02283912
- 417 Chéron B, Cronin AL, Doums C, et al (2011a) Unequal resource allocation among colonies produced  
418 by fission in the ant *Cataglyphis cursor*. *Ecology* 92:1448–1458. doi: 10.1890/10-2347.1
- 419 Chéron B, Monnin T, Fédérici P, Doums C (2011b) Variation in patriline reproductive success during  
420 queen production in orphaned colonies of the thelytokous ant *Cataglyphis cursor*. *Mol Ecol*  
421 20:2011–2022. doi: 10.1111/j.1365-294X.2011.05075.x
- 422 Clutton-Brock T, Mcauliffe K (2009) Female mate choice in mammals. *Q Rev Biol* 84:3–27.
- 423 Cronin AL, Fédérici P, Doums C, Monnin T (2012) The influence of intraspecific competition on  
424 resource allocation during dependent colony foundation in a social insect. *Oecologia* 168:361–

425 369. doi: 10.1007/s00442-011-2098-6

426 Cronin AL, Molet M, Doums C, et al (2013) Recurrent evolution of Dependent Colony Foundation  
427 across eusocial Insects. *Annu Rev Entomol* 58:120830113030002. doi: 10.1146/annurev-ento-  
428 120811-153643

429 Cronin AL, Monnin T, Haussy C, Doums C (2011) Opportunities for mate choice in the fission-  
430 performing ant *Cataglyphis cursor*. *Ecol Entomol* 36:522–525. doi: 10.1111/j.1365-  
431 2311.2011.01291.x

432 Dahbi A, Hefetz A, Lenoir A (2008) Chemotaxonomy of some *Cataglyphis* ants from Morocco and  
433 Burkina Faso. *Biochem Syst Ecol* 36:564–572. doi: 10.1016/j.bse.2008.03.004

434 Dahbi A, Lenoir A (1998) Queen and colony odour in the multiple nest ant species, *Cataglyphis iberica*  
435 (Hymenoptera, Formicidae). *Insectes Soc* 45:301–313. doi: 10.1007/s000400050090

436 de Souza EA, Trigo JR, Santos DE, et al (2017) The relationship between queen execution and  
437 cuticular hydrocarbons in stingless bee *Melipona scutellaris* (Hymenoptera: Meliponini).  
438 *Chemoecology* 27:25–32. doi: 10.1007/s00049-016-0226-9

439 Doums C, Cronin AL, Ruel C, et al (2013a) Facultative use of thelytokous parthenogenesis for queen  
440 production in the polyandrous ant *Cataglyphis cursor*. *J Evol Biol* 26:1431–1444. doi:  
441 10.1111/jeb.12142

442 Doums C, Ruel C, Clémencet J, et al (2013b) Fertile diploid males in the ant *Cataglyphis cursor*: A  
443 potential cost of thelytoky? *Behav Ecol Sociobiol* 67:1983–1993. doi: 10.1007/s00265-013-  
444 1606-6

445 Evison SEF, Ferreira RS, D’Ettorre P, et al (2012) Chemical signature and reproductive status in the  
446 facultatively polygynous ant *Pachycondyla verena*. *J Chem Ecol* 38:1441–1449. doi:  
447 10.1007/s10886-012-0195-6

448 Foitzik S, Fröba J, Rüger MH, Witte V (2011) Competition over workers: Fertility signalling in wingless  
449 queens of *Hypoconerops opacior*. *Insectes Soc* 58:271–278. doi: 10.1007/s00040-011-0147-0

450 Genuer R, Poggi J-M, Tuleau-Malot C (2015) VSURF: An R package for variable selection using random

451 forests. R J 7:19–33.

452 Gökçen OA, Morgan ED, Dani FR, et al (2002) Dufour gland contents of ants of the *Cataglyphis bicolor*  
453 group. J Chem Ecol 28:71–87. doi: 10.1023/A:1013521123356

454 Goldstein BA, Polley EC, Briggs FBS (2011) Random forests for genetic association studies. Stat Appl  
455 Genet Mol Biol 10:32. doi: 10.2202/1544-6115.1691

456 Greenberg L, Tröger AG, Francke W, et al (2007) Queen sex pheromone of the slave-making ant,  
457 *Polyergus breviceps*. J Chem Ecol 33:935–945. doi: 10.1007/s10886-007-9269-2

458 Hannonen M, Sledge MF, Turillazzi S, et al (2002) Queen reproduction, chemical signalling and  
459 worker behaviour in polygyne colonies of the ant *Formica fusca*. Anim Behav 64:477–485. doi:  
460 10.1006/anbe.2002.4001

461 Hefetz A (2007) The evolution of hydrocarbon pheromone parsimony in ants (Hymenoptera:  
462 Formicidae) – interplay of colony odor uniformity and odor idiosyncrasy. A review.  
463 Myrmecological News 10:59–68.

464 Hefetz A, Lenoir A (1992) Dufour’s gland composition in the desert ant *Cataglyphis* - species  
465 specificity and population differences. Zeitschrift Fur Naturforsch C-a J Biosci 47:285–289.

466 Heinze J, d’Ettorre P (2009) Honest and dishonest communication in social Hymenoptera. J Exp Biol  
467 212:1775–1779. doi: 10.1242/jeb.015008

468 Helft F, Doums C, Monnin T (2016) No evidence of pre-copulatory mate choice by gynes in the  
469 facultatively parthenogenetic ant *Cataglyphis cursor*. Insectes Soc 63:199–201. doi:  
470 10.1007/s00040-015-0444-0

471 Helft F, Monnin T, Doums C (2015) First evidence of inclusive sexual selection in the ant *Cataglyphis*  
472 *cursor*: worker aggressions differentially affect male access to virgin queens. Ethology 121:641–  
473 650.

474 Hölldobler B, Wilson EO (1990) The ants. Springer Verlag

475 Holman L, Jørgensen CG, Nielsen J, d’Ettorre P (2010) Identification of an ant queen pheromone  
476 regulating worker sterility. Proc R Soc London B 277:3793–3800. doi: 10.1098/rspb.2010.0984

477 Holman L, Lanfear R, D’Ettorre P (2013) The evolution of queen pheromones in the ant genus *Lasius*.  
478 J Evol Biol 26:1549–1558. doi: 10.1111/jeb.12162

479 Hora RR, Ionescu-Hirsh A, Simon T, et al (2008) Postmating changes in cuticular chemistry and visual  
480 appearance in *Ectatomma tuberculatum* queens (Formicidae: Ectatomminae).  
481 Naturwissenschaften 95:55–60. doi: 10.1007/s00114-007-0287-2

482 Jansen J, Pokorny T, Schmitt T (2016) Disentangling the effect of insemination and ovary  
483 development on the cuticular hydrocarbon profile in the bumblebee *Bombus terrestris*  
484 (Hymenoptera: Apidae). Apidologie 47:101–113. doi: 10.1007/s13592-015-0379-5

485 Jennions MD, Petrie M (1997) Variation in mate choice and mating preferences: A review of causes  
486 and consequences. Biol Rev 72:283–327. doi: 10.1017/s0006323196005014

487 Keegans SJ, Morgan ED, Agosti D, Wehner R (1992) What do glands tell us about species? A chemical  
488 case study of *Cataglyphis* ants. Biochem Syst Ecol 20:559–572. doi: 10.1016/0305-  
489 1978(92)90010-B

490 Keller L, Nonacs P (1993) The role of queen pheromones in social insects: queen control or queen  
491 signal? Anim. Behav. 45:787–794.

492 Le Conte Y, Hefetz A (2008) Primer pheromones in social Hymenoptera. Annu Rev Entomol 53:523–  
493 542. doi: 10.1146/annurev.ento.52.110405.091434

494 Lenoir A, Quérard L, Pondicq N, Berton F (1988) Reproduction and dispersal in the ant *Cataglyphis*  
495 *cursor* (Hymenoptera, Formicidae). Psyche 95:21–44. doi: 10.1155/1988/54685

496 Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2:18–22. doi:  
497 10.1177/154405910408300516

498 Liu J, Dong M, Miao SL, et al (2006) Invasive alien plants in China: Role of clonality and geographical  
499 origin. Biol Invasions 8:1461–1470. doi: 10.1007/s10530-005-5838-x

500 Mayade S, Cammaerts M-C, Suzzoni J-P (1993) Home-range marking and territorial marking in  
501 *Cataglyphis cursor* (Hymenoptera, Formicidae). Behav Processes 30:131–142. doi:  
502 10.1016/0376-6357(93)90003-A

503 Mitra A, Ramachandran A, Gadagkar R (2014) Nestmate discrimination in the social wasp *Ropalidia*  
504 *marginata*: chemical cues and chemosensory mechanism. *Anim Behav* 88:113–124. doi:  
505 10.1016/j.anbehav.2013.11.017

506 Monnin T (2006) Chemical recognition of reproductive status in social insects. *Ann Zool Fennici*  
507 43:515–530.

508 Monnin T, Ratnieks FLW, Jones GR, Beard R (2002) Pretender punishment induced by chemical  
509 signalling in a queenless ant. *Nature* 419:61–65.

510 Nowbahari E, Lenoir A, Clément J-L, et al (1990) Individual, geographical and experimental variation  
511 of cuticular hydrocarbons of the ant *cataglyphis cursor* (Hymenoptera: Formicidae): Their use in  
512 nest and subspecies recognition. *Biochem Syst Ecol* 18:63–73. doi: 10.1016/0305-  
513 1978(90)90036-F

514 Oi CA, Millar JG, van Zweden JS, Wenseleers T (2016) Conservation of queen pheromones across two  
515 species of vespine wasps. *J Chem Ecol* 42:1175–1180. doi: 10.1007/s10886-016-0777-9

516 Oi CA, van Zweden JS, Oliveira RC, et al (2015) The origin and evolution of social insect queen  
517 pheromones: novel hypotheses and outstanding problems. *BioEssays* 37:808–821. doi:  
518 10.1002/bies.201400180

519 Oliveira RC, Oi CA, Vollet-Neto A, Wenseleers T (2016) Intraspecific worker parasitism in the common  
520 wasp, *Vespula vulgaris*. *Anim Behav* 113:79–85. doi: 10.1016/j.anbehav.2015.12.025

521 Oppelt A, Heinze J (2009) Mating is associated with immediate changes of the hydrocarbon profile of  
522 *Leptothorax gredleri* ant queens. *J Insect Physiol* 55:624–628. doi:  
523 10.1016/j.jinsphys.2009.03.010

524 R Core Team (2016) R: A language and environment for statistical computing.

525 Ranganathan Y, Borges RM (2010) Reducing the babel in plant volatile communication: using the  
526 forest to see the trees. *Plant Biol* 12:735–742. doi: 10.1111/j.1438-8677.2009.00278.x

527 RStudio Team (2016) RStudio: Integrated development for R.

528 Ruel C, Lenoir A, Cerdá X, Boulay R (2013) Surface lipids of queen-laid eggs do not regulate queen

529 production in a fission-performing ant. *Naturwissenschaften* 100:91–100. doi: 10.1007/s00114-  
530 012-0997-y

531 Smith AA, Hölldobler B, Liebig J (2012) Queen-specific signals and worker punishment in the ant  
532 *Aphaenogaster cockerelli*: the role of the Dufour’s gland. *Anim Behav* 83:587–593. doi:  
533 10.1016/j.anbehav.2011.12.024

534 Smith AA, Millar JG, Suarez A V. (2016) Comparative analysis of fertility signals and sex-specific  
535 cuticular chemical profiles of *Odontomachus* trap-jaw ants. *J Exp Biol* 219:419–30. doi:  
536 10.1242/jeb.128850

537 Stearns SC, Hoekstra R (2005) *Evolution. An introduction. Second edition.* Oxford University Press

538 van Oystaeyen A, Oliveira RC, Holman L, et al (2014) Conserved class of queen pheromones stops  
539 social insect workers from reproducing. *Science* 287:287–291. doi: 10.1126/science.1244899

540 van Zweden JS, Bonckaert W, Wenseleers T, D’Ettorre P (2014) Queen signaling in social wasps.  
541 *Evolution* 68:976–986. doi: 10.1111/evo.12314

542 van Zweden JS, D’Ettorre P (2010) The role of hydrocarbons in nestmate recognition. In: Blomquist  
543 GJ, Bagnères A-G (eds) *Insect Hydrocarbons: Biology, Biochemistry and Chemical Ecology*,  
544 Cambridge. Cambridge, pp 222–243

545 Walter F, Fletcher DJC, Chautems D, et al (1993) Identification of the sex pheromone of an ant,  
546 *Formica lugubris* (Hymenoptera, Formicidae). *Naturwissenschaften* 80:30–34. doi:  
547 10.1007/BF01139755

548

549 **Fig. 1** MDS plots of the proximity distance estimated from the Random Forest analysis including the  
550 four groups of individuals. The first two dimensions of the multidimensional distance matrix are  
551 shown for cuticular hydrocarbons and Dufour's gland content.

552

553

554 **Fig. 2** Random forest analyses for cuticular hydrocarbons, comparing queens and workers (a), gynes  
555 G1 and gynes G0 (b), and gynes G1 and queens (c). For clarity, each plot only shows the 20 peaks  
556 with the highest averaged rank of importance. The plot shows average ranks, as well as minimal and  
557 maximal ranks when some variation occurred. The colour of the dots indicates the group of female in  
558 which the peak was more abundant (values from Table S1). Peaks importance for discriminating the  
559 two groups of females decrease from rank 1, the most important peak, to rank 20, the least  
560 important peak. Peaks needed to obtain the best discrimination (i.e. retained by the second selection  
561 step, see M&M) are encircled. Peak identification is given in Table S1.

562

563

564 **Fig. 3** Random forest analyses for Dufour's gland content, comparing queens and workers (a), gynes  
565 G1 and gynes G0 (b), and gynes G1 and queens (c). Same legend than Figure 2. Peak identification  
566 and abundance are given in Table S2.

567



568 **Table 1** Colony origin of the 54 individuals analysed. Three colonies were queenless (QL) at the time  
 569 of gyne production, and two colonies produced males in addition to gynes  
 570

COLONY CODE	COLONY SIZE	MALES PRODUCED	INDIVIDUALS SAMPLED			
			QUEEN	GYNES G0	GYNES G1	WORKER
1206 QL*	606			2		1
1250	254		1			1
1251	1561		1			1
1252	877		1			1
1254	217		1			1
1257	430		1			1
1258 QL	459	4		5	6	1
1259 QL	1857			2	4	1
1260	168		1			1
1262	867		1			1
1263	299		1			1
1264	912		1			1
1265	185		1			1
1266	361		1			1
1267	883		1			1
1268	340	23	1	3	3	1
TOTAL			13	12	13	16

571 \* collected in March 2014

572

Fig. 1

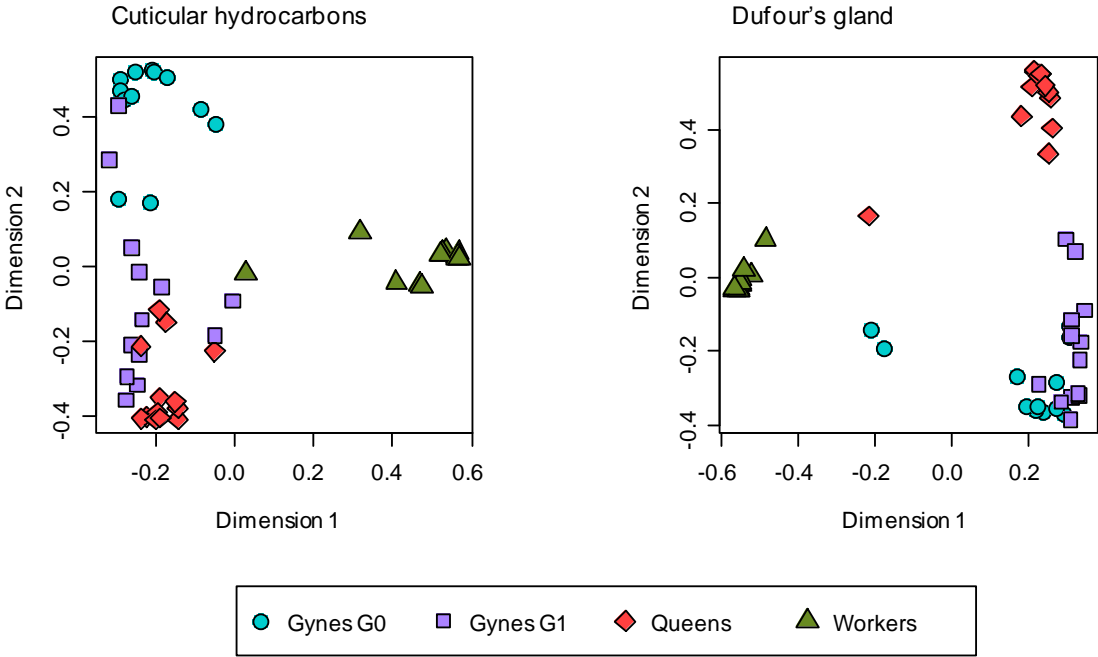


Fig. 2

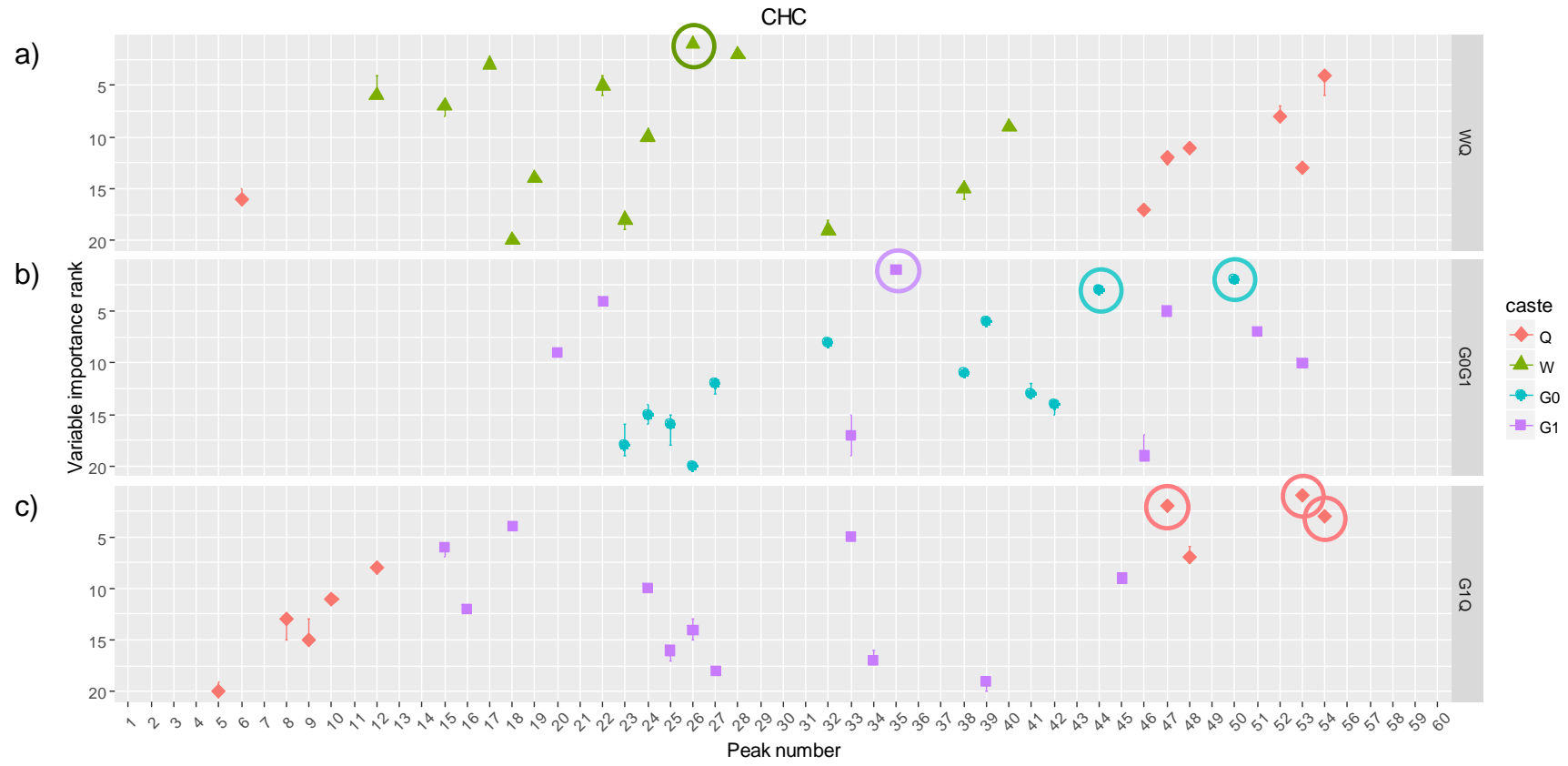
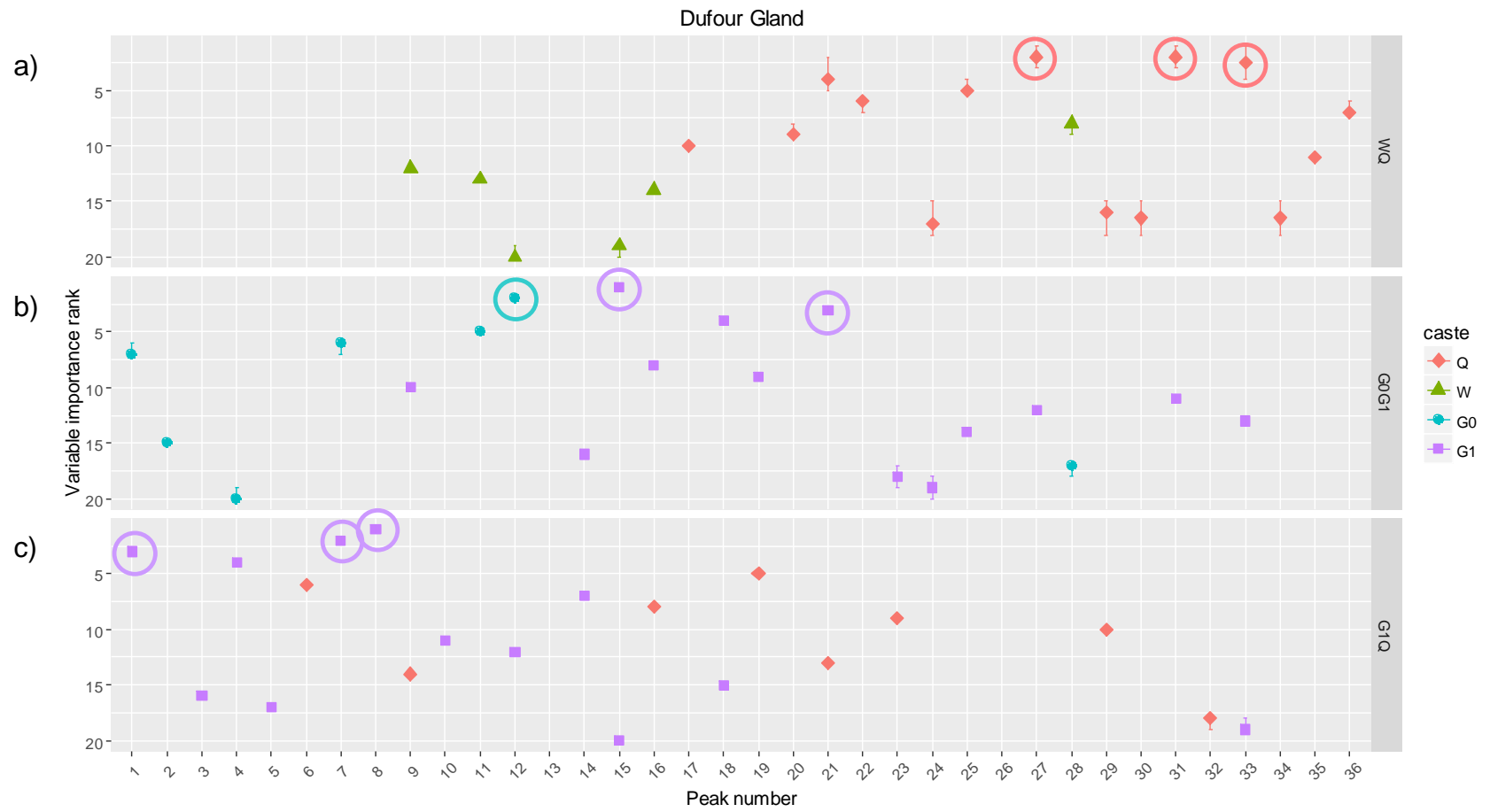


Fig. 3



## Supplementary materials

**Table S1** Relative abundance of peaks (percentage of total area) in extracts of cuticular hydrocarbons (CHC) of workers, gynes G0 and G1, and queens

PEAK NUMBER	PEAK IDENTIFICATION	RETENTION TIME	PEAK RELATIVE ABUNDANCE (MEAN % ± SD)			
			WORKERS	GYNES G0	GYNES G1	QUEENS
P1	C <sub>21:1</sub>	19.36	0.12 ± 0.44	0.04 ± 0.07	0.18 ± 0.41	0.58 ± 0.96
P2	n-C <sub>21</sub>	19.78	0.06 ± 0.04	0.05 ± 0.09	0.15 ± 0.26	0.35 ± 0.61
P3	x,y-dimeC <sub>21</sub>	20.40	1.01 ± 0.44	0.17 ± 0.10	0.41 ± 0.34	1.43 ± 2.05
P4	x,y-dimeC <sub>21</sub>	21.06	2.01 ± 0.82	0.35 ± 0.18	0.65 ± 0.31	1.27 ± 0.70
P5	n-C <sub>22</sub>	21.37	0.31 ± 0.12	0.03 ± 0.05	0.04 ± 0.07	0.15 ± 0.12
P6	C <sub>23:1</sub>	22.56	0.01 ± 0.03	0.09 ± 0.14	0.42 ± 1.07	1.05 ± 2.01
P7	n-C <sub>23</sub>	22.97	0.14 ± 0.13	0.14 ± 0.18	0.32 ± 0.63	0.74 ± 1.60
P8	x,y-dimeC <sub>23</sub>	23.77	0.62 ± 0.29	0.10 ± 0.07	0.20 ± 0.09	0.37 ± 0.18
P9	x,y-dimeC <sub>23</sub>	24.41	1.35 ± 0.61	0.21 ± 0.14	0.43 ± 0.18	0.85 ± 0.49
P10	n-C <sub>24</sub>	24.60	0.73 ± 0.23	0.12 ± 0.08	0.16 ± 0.10	0.57 ± 0.81
P11	C <sub>25:1</sub>	25.65	1.16 ± 2.41	0.97 ± 0.26	1.68 ± 1.48	1.71 ± 3.33
P12	n-C <sub>25</sub>	26.04	9.22 ± 6.01	1.57 ± 0.33	2.27 ± 1.06	1.17 ± 0.28
P13	11+13-meC <sub>25</sub>	26.54	0.51 ± 0.23	0.42 ± 0.06	0.37 ± 0.10	0.34 ± 0.16
P14	5-meC <sub>25</sub>	26.79	0.42 ± 0.24	0.30 ± 0.08	0.28 ± 0.14	0.26 ± 0.14
P15	3-meC <sub>25</sub>	27.11	2.04 ± 0.90	1.51 ± 0.46	1.30 ± 0.43	0.69 ± 0.19
P16	5,9-dimeC <sub>25</sub>	27.25	0.29 ± 0.33	0.39 ± 0.07	0.29 ± 0.10	0.12 ± 0.08
P17	n-C <sub>26</sub>	27.51	1.50 ± 0.63	0.19 ± 0.11	0.27 ± 0.09	0.22 ± 0.19
P18	3,7-dimeC <sub>25</sub>	27.63	0.98 ± 0.49	1.57 ± 0.60	0.98 ± 0.34	0.41 ± 0.23
P19	4-meC <sub>26</sub>	28.34	0.69 ± 0.19	0.64 ± 0.22	0.48 ± 0.20	0.26 ± 0.15
P20	C <sub>27:1</sub>	28.60	2.57 ± 1.45	3.55 ± 1.15	6.97 ± 2.15	4.44 ± 2.60
P21	4,8-dimeC <sub>26</sub>	28.81	1.54 ± 3.47	1.48 ± 0.40	1.06 ± 0.43	0.58 ± 0.21
P22	n-C <sub>27</sub>	28.94	10.35 ± 4.07	1.48 ± 0.31	3.58 ± 1.11	2.29 ± 0.74
P23	11+13-meC <sub>27</sub>	29.40	2.68 ± 0.40	4.09 ± 0.43	2.83 ± 0.98	1.81 ± 0.72
P24	5-meC <sub>27</sub>	29.65	4.43 ± 1.60	3.93 ± 1.10	2.85 ± 1.01	1.43 ± 0.65
P25	11,15-dimeC <sub>27</sub>	29.80	0.40 ± 0.74	0.58 ± 0.18	0.39 ± 0.19	0.13 ± 0.14
P26	3-meC <sub>27</sub>	29.97	3.68 ± 0.92	2.09 ± 0.69	1.67 ± 0.45	0.89 ± 0.46
P27	5,9-dimeC <sub>27</sub>	30.08	3.43 ± 1.17	5.47 ± 1.16	3.82 ± 1.32	2.12 ± 0.99
P28	n-C <sub>28</sub>	30.34	1.04 ± 0.77	0.62 ± 1.28	0.20 ± 0.09	0.17 ± 0.12

P29	3,11-dimeC <sub>27</sub>	30.48	4.39 ± 1.71	8.45 ± 3.28	6.00 ± 2.54	3.40 ± 1.50
P30	12-meC <sub>28</sub>	30.77	0.79 ± 0.34	1.47 ± 0.36	0.92 ± 0.35	0.63 ± 0.31
P31	8,12-dimeC <sub>28</sub>	30.82	0.96 ± 0.46	2.40 ± 0.74	1.60 ± 0.80	0.85 ± 0.65
P32	4-meC <sub>28</sub>	31.13	0.58 ± 0.23	0.40 ± 0.06	0.25 ± 0.12	0.19 ± 0.19
P33	C <sub>29:1</sub>	31.37	1.75 ± 0.76	2.44 ± 0.54	3.40 ± 0.77	2.23 ± 1.06
P34	4,12-dimeC <sub>28</sub>	31.53	1.70 ± 1.63	3.20 ± 0.67	2.26 ± 0.64	1.28 ± 0.55
P35	<i>n</i> -C <sub>29</sub>	31.68	5.55 ± 1.16	0.75 ± 0.27	3.18 ± 1.04	4.61 ± 1.55
P36	11+13+15-meC <sub>29</sub>	32.09	3.45 ± 1.08	7.28 ± 2.25	5.45 ± 1.60	5.08 ± 1.98
P37	7-meC <sub>29</sub>	32.21	0.56 ± 0.20	0.77 ± 0.11	0.59 ± 0.25	0.46 ± 0.24
P38	5-meC <sub>29</sub>	32.34	1.32 ± 0.44	0.84 ± 0.22	0.65 ± 0.17	0.47 ± 0.20
P39	7,11-dimeC <sub>29</sub>	32.60	0.36 ± 0.19	0.81 ± 0.14	0.43 ± 0.16	0.20 ± 0.19
P40	3-meC <sub>29</sub>	32.64	0.89 ± 0.47	0.40 ± 0.19	0.35 ± 0.09	0.22 ± 0.18
P41	5,9-dimeC <sub>29</sub>	32.74	2.40 ± 0.84	4.26 ± 0.62	2.75 ± 0.99	1.67 ± 0.87
P42	3,13-dimeC <sub>29</sub>	33.04	2.15 ± 0.79	4.50 ± 0.88	2.64 ± 1.08	1.65 ± 0.90
P43	3,7-dimeC <sub>29</sub>	33.09	1.17 ± 0.39	1.36 ± 0.27	0.10 ± 0.43	0.71 ± 0.25
P44	14-meC <sub>30</sub> + 8,12-dimeC <sub>30</sub>	33.44	1.91 ± 0.77	3.59 ± 0.42	2.06 ± 0.58	1.80 ± 0.61
P45	C <sub>31:2</sub>	33.77	0.82 ± 0.35	1.06 ± 0.40	0.85 ± 0.38	0.52 ± 0.16
P46	C <sub>31:1</sub>	34.00	1.55 ± 0.58	2.65 ± 0.79	4.70 ± 1.66	4.47 ± 2.00
P47	<i>n</i> -C <sub>31</sub>	34.26	2.10 ± 0.83	0.20 ± 0.13	1.82 ± 1.09	7.62 ± 2.86
P48	13+15-meC <sub>31</sub>	34.62	2.27 ± 0.98	4.92 ± 2.68	4.17 ± 0.74	7.12 ± 3.22
P49	7-meC <sub>31</sub>	34.76	0.50 ± 0.43	0.29 ± 0.11	0.40 ± 0.12	0.55 ± 0.34
P50	11,13-dimeC <sub>31</sub>	35.55	1.65 ± 0.53	2.42 ± 0.37	1.27 ± 0.41	0.99 ± 0.74
P51	C <sub>33:2</sub>	36.17	1.54 ± 1.30	3.11 ± 2.81	6.62 ± 3.97	6.93 ± 3.84
P52	C <sub>33:1</sub>	36.42	2.56 ± 1.24	5.00 ± 1.90	6.42 ± 1.45	8.84 ± 2.65
P53	<i>n</i> -C <sub>33</sub>	36.68	0.57 ± 0.59	0.07 ± 0.14	1.00 ± 0.59	4.52 ± 2.03
P54	15+17-meC <sub>33</sub>	37.00	0.45 ± 0.25	1.15 ± 0.83	0.92 ± 0.31	1.90 ± 0.79
P55	C <sub>34:1</sub>	37.44	0.07 ± 0.11	0.42 ± 0.28	0.24 ± 0.29	0.22 ± 0.52
P56	4, <i>y</i> -dimeC <sub>33</sub>	37.98	0.50 ± 1.15	0.40 ± 0.18	0.32 ± 0.47	0.33 ± 0.80
P57	C <sub>35:2</sub>	38.45	0.55 ± 0.38	0.65 ± 0.48	1.26 ± 1.27	1.41 ± 1.00
P58	C <sub>35:1</sub>	38.67	1.36 ± 0.46	1.99 ± 0.62	2.03 ± 1.46	2.58 ± 2.03
P59	<i>x,y</i> -dimeC <sub>35</sub>	40.12	0.28 ± 0.76	0.24 ± 0.17	0.07 ± 0.15	0.07 ± 0.18
P60	C <sub>37:1</sub>	40.77	0	0.27 ± 0.28	0.18 ± 0.21	0.04 ± 0.13

**Table S2** Relative abundance (percentage of total area) of peaks in extracts of Dufour's gland (DG) of workers, gynes G0 and G1, and queens

PEAK NUMBER	PEAK IDENTIFICATION	RETENTION TIME	PEAK RELATIVE ABUNDANCE (MEAN % $\pm$ SD)			
			WORKERS	GYNES G0	GYNES G1	QUEENS
P1	<i>n</i> -C <sub>11</sub>	5.72	0.08 $\pm$ 0.22	5.39 $\pm$ 2.63	2.79 $\pm$ 0.94	0.64 $\pm$ 0.49
P2	<i>n</i> -C <sub>12</sub>	7.15	0.01 $\pm$ 0.02	0.26 $\pm$ 0.18	0.13 $\pm$ 0.05	0.11 $\pm$ 0.07
P3	<i>n</i> -C <sub>13</sub>	8.54	4.53 $\pm$ 3.80	15.84 $\pm$ 6.29	12.24 $\pm$ 3.42	10.41 $\pm$ 3.48
P4	Dodecyl acetate	9.98	0.02 $\pm$ 0.03	0.29 $\pm$ 0.24	0.23 $\pm$ 0.14	0.00 $\pm$ 0.00
P5	Dodecanol	10.82	0.27 $\pm$ 0.43	0.93 $\pm$ 1.45	0.31 $\pm$ 0.35	0.09 $\pm$ 0.13
P6	<i>n</i> -C <sub>15</sub>	11.09	2.56 $\pm$ 1.68	2.13 $\pm$ 1.32	2.10 $\pm$ 0.42	3.24 $\pm$ 0.67
P7	Dodecanyl acetate	12.46	21.76 $\pm$ 8.46	20.72 $\pm$ 5.23	14.41 $\pm$ 3.63	5.21 $\pm$ 3.32
P8	Tridecanol	12.78	0.20 $\pm$ 0.20	0.36 $\pm$ 0.23	0.41 $\pm$ 0.10	0.08 $\pm$ 0.06
P9	<i>n</i> -C <sub>17</sub>	13.67	6.78 $\pm$ 1.23	1.49 $\pm$ 0.96	2.02 $\pm$ 0.36	2.69 $\pm$ 0.95
P10	? (fatty acid)	13.80	0.32 $\pm$ 0.16	0.89 $\pm$ 0.47	1.10 $\pm$ 0.19	0.76 $\pm$ 0.32
P11	meC <sub>17</sub>	13.84	0.72 $\pm$ 1.13	0.97 $\pm$ 2.03	0.05 $\pm$ 0.05	0.06 $\pm$ 0.04
P12	meC <sub>17</sub>	14.44	1.29 $\pm$ 2.05	1.77 $\pm$ 3.46	0.21 $\pm$ 0.31	0.09 $\pm$ 0.07
P13	Tetradecanol acetate	15.22	1.42 $\pm$ 0.66	3.25 $\pm$ 2.21	4.21 $\pm$ 1.35	4.29 $\pm$ 2.22
P14	Farnesol acetate	15.75	0	0.14 $\pm$ 0.15	0.26 $\pm$ 0.11	0.06 $\pm$ 0.06
P15	C <sub>19:2</sub>	16.14	2.90 $\pm$ 0.98	0.19 $\pm$ 0.18	1.10 $\pm$ 0.82	0.93 $\pm$ 0.43
P16	C <sub>19:1</sub>	16.25	44.39 $\pm$ 9.89	4.90 $\pm$ 2.83	9.86 $\pm$ 3.10	17.62 $\pm$ 6.49
P17	<i>n</i> -C <sub>19</sub>	16.60	1.32 $\pm$ 0.31	2.92 $\pm$ 1.80	3.71 $\pm$ 1.04	3.59 $\pm$ 1.01
P18	C <sub>21:2</sub>	19.27	1.20 $\pm$ 0.55	0.51 $\pm$ 0.49	1.94 $\pm$ 1.34	1.88 $\pm$ 0.52
P19	C <sub>21:1</sub>	19.36	2.76 $\pm$ 0.96	1.17 $\pm$ 0.75	2.44 $\pm$ 0.80	4.41 $\pm$ 0.87
P20	<i>n</i> -C <sub>21</sub>	19.77	0.07 $\pm$ 0.11	2.41 $\pm$ 1.48	2.87 $\pm$ 1.33	2.13 $\pm$ 0.70
P21	C <sub>23:1</sub>	22.57	0.23 $\pm$ 0.21	4.30 $\pm$ 1.11	5.29 $\pm$ 0.71	6.15 $\pm$ 1.54
P22	<i>n</i> -C <sub>23</sub>	23.00	0	4.85 $\pm$ 2.98	5.31 $\pm$ 2.05	4.21 $\pm$ 1.41
P23	Eicosenol	24.29	2.67 $\pm$ 0.89	0.91 $\pm$ 1.02	2.37 $\pm$ 1.45	5.52 $\pm$ 1.83
P24	Eicosyl acetate	24.68	0.02 $\pm$ 0.05	0.38 $\pm$ 0.37	0.56 $\pm$ 0.25	0.83 $\pm$ 0.43
P25	C <sub>25:1</sub>	25.81	0	6.39 $\pm$ 1.79	7.56 $\pm$ 1.00	7.98 $\pm$ 2.26
P26	<i>n</i> -C <sub>25</sub>	26.03	0.08 $\pm$ 0.13	0.59 $\pm$ 0.53	0.48 $\pm$ 0.35	0.42 $\pm$ 0.29
P27	C <sub>27:1</sub>	28.60	0.01 $\pm$ 0.04	3.24 $\pm$ 1.68	4.41 $\pm$ 0.55	4.69 $\pm$ 0.80

P28	Dodecyl myristate	29.91	4.24 ± 2.578	0.89 ± 1.23	0.28 ± 0.28	0.31 ± 0.27
P29	Tetracosanal	30.22	0	0.32 ± 0.23	0.53 ± 0.18	0.83 ± 0.32
P30	Tetracosyl acetate	30.55	0	0.60 ± 0.54	0.86 ± 0.41	0.74 ± 0.38
P31	C <sub>29:1</sub>	31.45	0	3.36 ± 1.98	4.27 ± 1.09	3.56 ± 0.95
P32	C <sub>30</sub>	32.96	0.09 ± 0.24	0.28 ± 0.25	0.49 ± 0.25	0.74 ± 0.43
P33	C <sub>31:2</sub>	34.00	0	0.51 ± 0.37	1.00 ± 0.54	0.62 ± 0.24
P34	?	35.44	0	0.40 ± 0.30	0.63 ± 0.25	0.93 ± 0.45
P35	?	37.79	0.06 ± 0.25	1.81 ± 0.88	2.07 ± 0.65	2.66 ± 1.44
P36	?	40.11	0	4.60 ± 11.07	1.48 ± 0.51	1.51 ± 0.95



**Table S3** confusion matrices of the random forests performed with the four groups of individuals

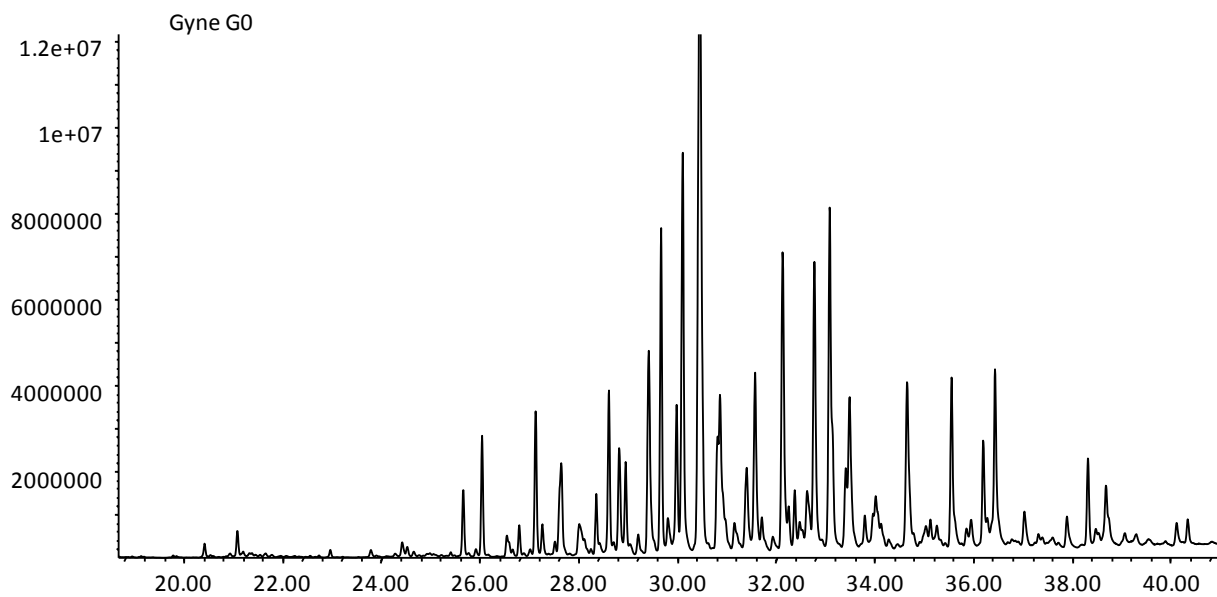
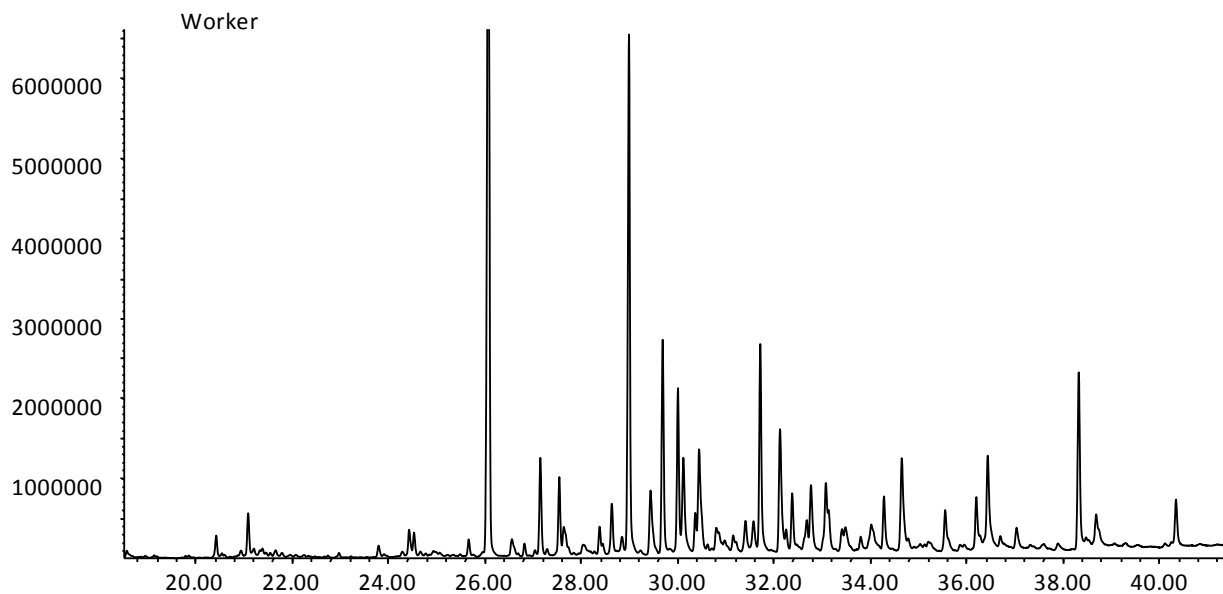
Cuticular hydrocarbons:

	G0	G1	Q	W	Classification error
G0	12	0	0	0	0
G1	2	10	1	0	0.231
Q	0	3	10	0	0.231
W	0	1	0	15	0.062

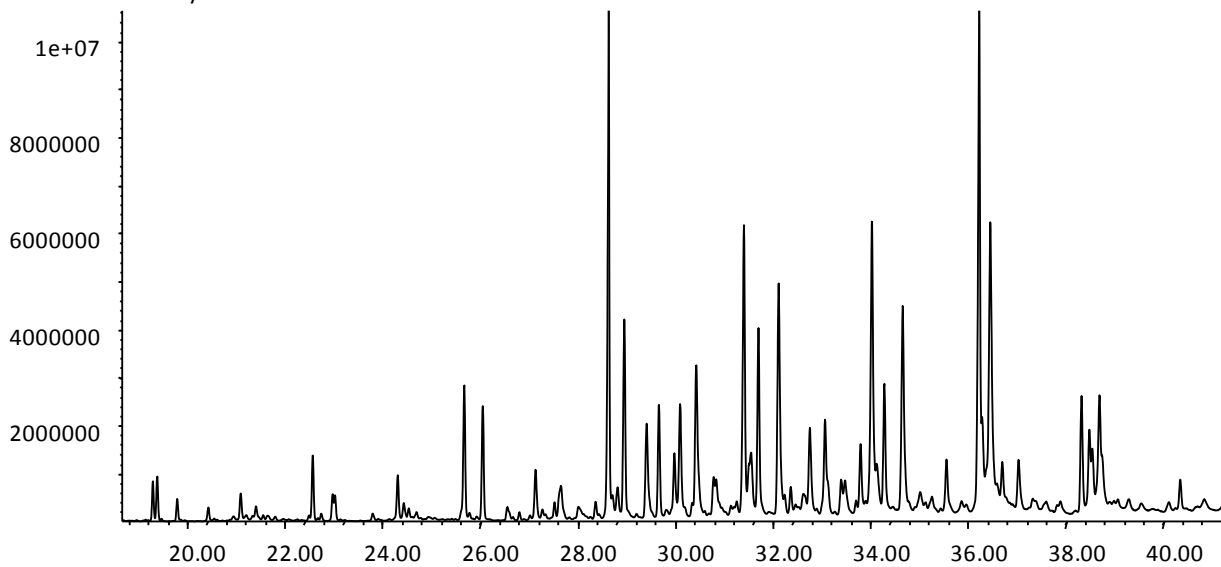
Dufour's gland:

	G0	G1	Q	W	Classification error
G0	9	3	0	0	0.250
G1	2	11	0	0	0.154
Q	0	0	12	1	0.077
W	0	0	0	16	0

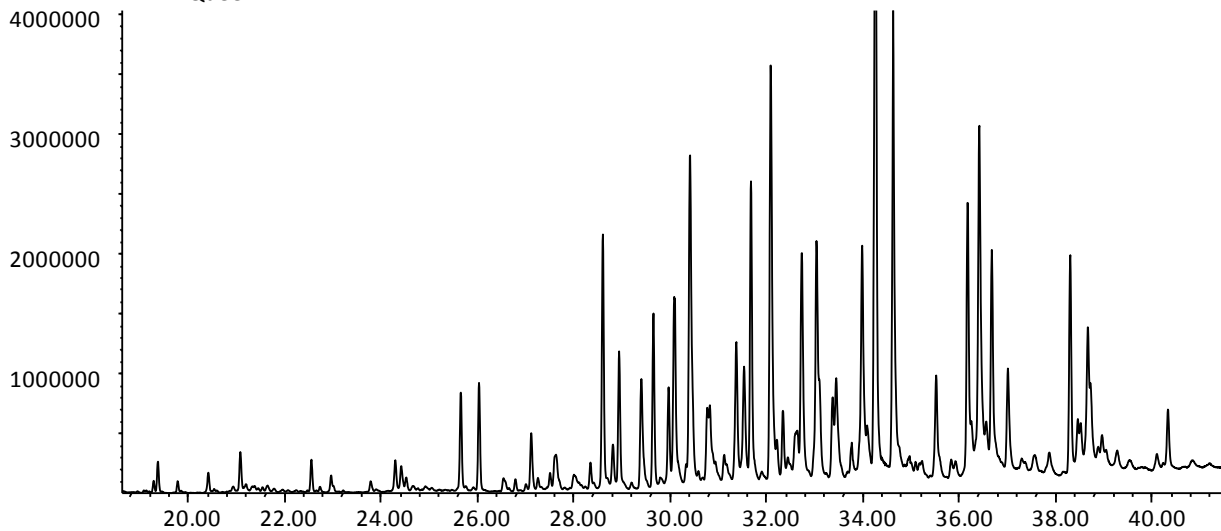
**Fig. S1** Representative chromatograms of cuticular hydrocarbon extracts for the four groups of individuals. X and y-axes are retention time and abundance, respectively.



Gyne G1



Queen



**Fig. S2** Representative chromatograms of Dufour's gland extracts for the four groups of individuals. X and y-axes are retention time and abundance, respectively.

