Supplemental Information

for

Epicuticular compounds of *Protopiophila litigata* (Diptera: Piophilidae): identification and sexual selection across two years in the wild

Annals of the Entomological Society of America

Christopher S. Angell[1], Sharon Curtis[2], Anaïs Ryckenbusch[2], and Howard D. Rundle[1]*

[1]*Department of Biology, University of Ottawa; Ottawa ON, K1N 6N5, Canada*

[2]*Department of Chemistry, University of Ottawa; Ottawa ON, K1N 6N5, Canada*

*Corresponding author, e-mail: hrundle@uottawa.ca.

Compound Identification

Gas chromatography with mass spectrometry (GC:MS) was used to generate mass spectra for individual epicuticular compounds (ECs), as described in "Methods and Materials" of the main text. Equivalent chain lengths (ECL) were also calculated for each chromatogram peak using a C_7-C_{40} alkane standard (Sigma Aldrich, product # 49452-U), following (Curtis et al. 2013). In general, MS fragmentation patterns and ECL numbers were cross referenced with databases, i.e. the Wiley 275 GC:MS database, NIST MS search 2.0, NIST Chemistry WebBook SRD 69 (https://webbook.nist.gov/chemistry), and Pherobase (www.pherobase.com) and published literature sources (McLafferty 1980, Lockey 1988, Howard and Blomquist 2005). ECL values themselves can also be used to make inferences about compound identities, as hydrocarbons with similar structures (e.g. locations of methyl groups or double bonds) exhibit the same fractional values (i.e. x.75), while differences in the integer portion reflect differences in overall carbon length (Bartelt et al. 1986).

Alkanes

Linear alkanes were identified by direct comparison of ECL and mass spectrum with the C_7-C_{40} alkane standard. For example, the ECL of heptacosane occurs at exactly at 27.00 and the mass spectrum is marked by a characteristic pattern of fragment ions, as can be observed in Fig. S1 (inset). FID peak 5 (see Table 1) of the antler fly sample also has a peak at the same ECL and retention time, with a nearly identical mass spectrum. This peak is therefore identified as heptocosane ($C_{27}H_{56}$, 380, Da). Similarly, FID peak 10 (ECL 29.00) was identified as nonacosane ($C_{29}H_{60}$, 408 Da) by comparison of the mass spectrum and ECL to nonacosane from the standard.

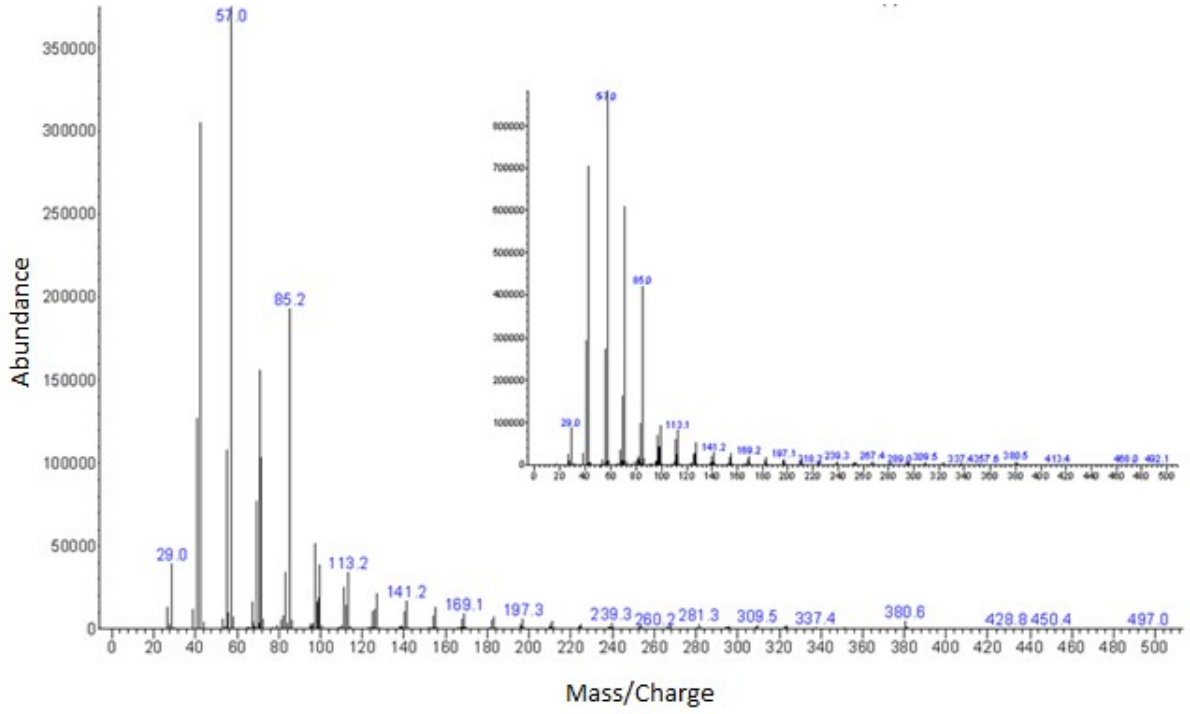

Fig. S1 Observed mass spectrum of FID peak 5 (heptacosane) from antler fly sample. The inset shows mass spectrum of heptacosane from the alkane standard.

Methyl Alkanes

In general, the mass spectra of methlyl alkanes are characterized by a stable secondary carbocation radical formed upon ionisation (McLafferty 1980). Fragmentation in these species tends to occur either before or after the methyl branch, and the charge is localised on the secondary cation, so these ions are represented in the mass spectrum. Therefore, the position of the side chain can be easily identified via the mass spectral fragmentation pattern. The identity of these compounds was also verified using the ECL and the Pherobase website (http://www.pherobase.com).

As an example, FID peak 14 has an ECL of 29.73, suggesting a hydrocarbon of length C_{29}-C_{30}. Fragmentation at the methyl side chain produced a stable secondary radical cation at mass 393.6 Da, corresponding to a C_{28} fragment (Fig. S2). Therefore, we identified the compound as 3-methylnonacosane ($C_{29}H_{60}$, 408 Da). Likewise, the mass spectrum of FID

peak 7 had an ECL of 27.73 and exhibited the presence of a secondary radical cation at mass 365.6, suggesting it is 3-methylheptacosane ($C_{28}H_{58}$, 394 Da). The fractional ECL value is the same as peak 14 (x.73), lending further support to the placement of the methyl group at the C_3 position. This same process was used to identify the remaining methyl alkanes.

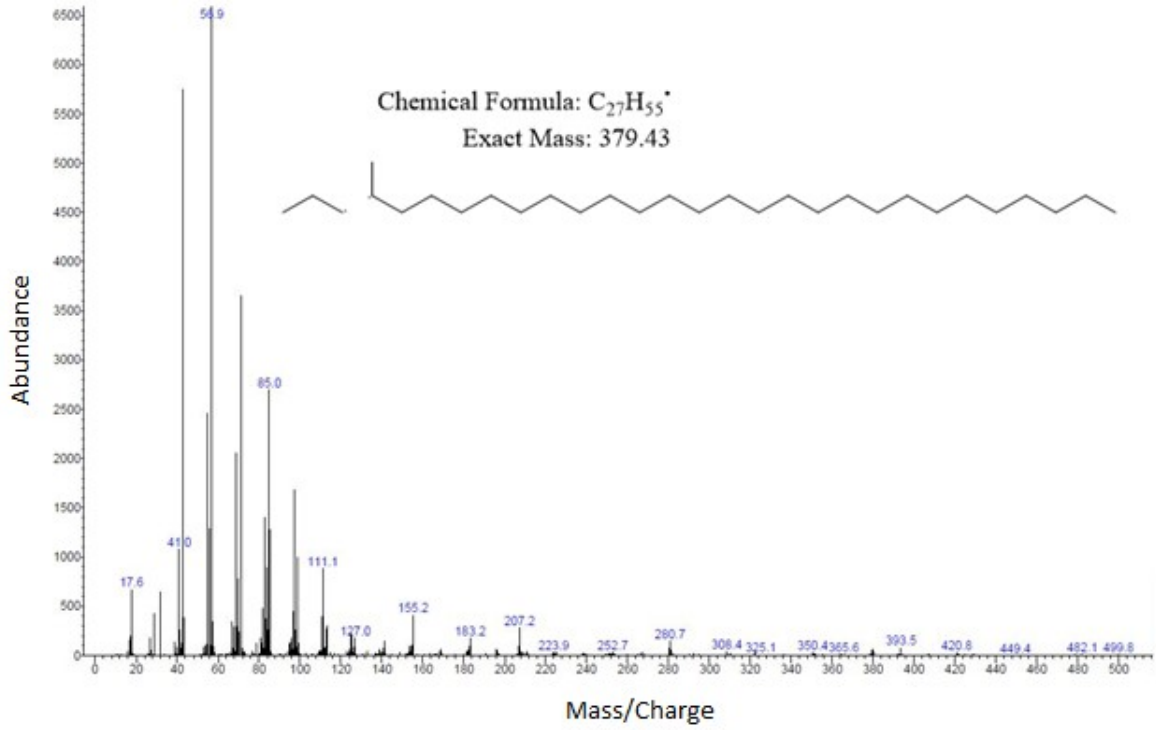

Fig. S2 Observed mass spectrum of FID peak 14 (3-methylnonacosane), including a stick diagram of the molecule to illustrate the fragmentation that produces the secondary radical cation.

Methyl Alkanes, coeluted

The spectra of peaks 6, 11 and 13 show many ions differing in mass by C_2H_4 due to the co-elution of two or three different methyl alkanes. The methyl branches are positioned near the centre of the molecule, so their precise location does not affect the boiling point, or the interaction with the column, and the compounds thus elute at the same time. The stability of the secondary radical cation produced by α cleavage at the branch allows for their

identification. The methyls are located near the centre of the chain, creating some symmetry by which the molecule produces two identical ions from each half of itself. Fig. S3 illustrates the origin of the ions visible in peak 6, a mixture of 11-methylheptacosane and 13-methylheptacosane. The other two peaks were identified in the same way.

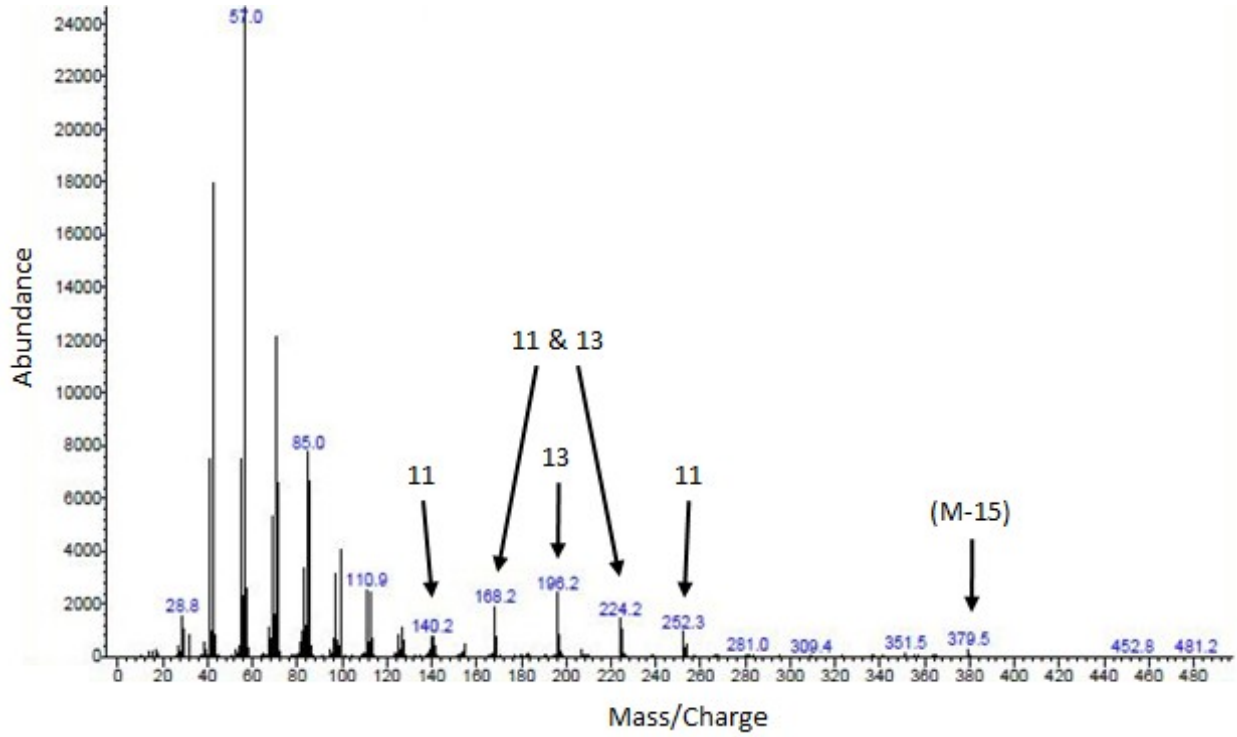

Fig. S3 Observed mass spectrum of FID peak 6, containing 11-methylheptocosane and 13-methylheptacosane, co-eluted. Numbers indicate which of the two species (or both) each diagnostic peak derives from.

Dimethyl Alkanes

Dimethyl alkanes can be identified in essentially the same way as methyl alkanes, as they also preferentially fragment adjacent to the methyl branches. Thus, for each methyl group there is a set of two peaks. The mass spectra are complicated by the fact that each of our dimethyl alkanes co-eluted with a second dimethyl species. Like many of the methyl alkanes above, the methyl branches are near the centre of the molecule, creating symmetry in

which the same size fragment can be produced in more than one way. In these cases, the ion often appears more intense on the mass spectrum. Using FID peak 13 (ECL 29.58) as an example, Table S1 demonstrates the masses produced by these cleavages and Fig. S4 illustrates the diagnostic ions used to confirm these chemical species.

Table S1 Diagnostic MS ion fragments caused by α cleavage in 11,15-dimethylnonacosane and 13,17-dimethylnonacosane, which co-elute as FID peak 13.

Location of cleavage	Ion fragments produced (m/z)
11,15-dimethylnonacosane	
11-methyl	141 (C_{10}) and 295 (C_{21}) 169 (C_{12}) and 267 (C_{19})
15-methyl	211 (C_{15}) and 225 (C_{16}) 197 (C_{14}) and 239 (C_{17})
13,17-dimethylnonacosane	
13-methyl	169 (C_{12}) and 267 (C_{19}) 197 (C_{14}) and 239 (C_{17})
17-methyl	239 (C_{17}) and 197 (C_{14}) 169 (C_{12}) and 267 (C_{19})

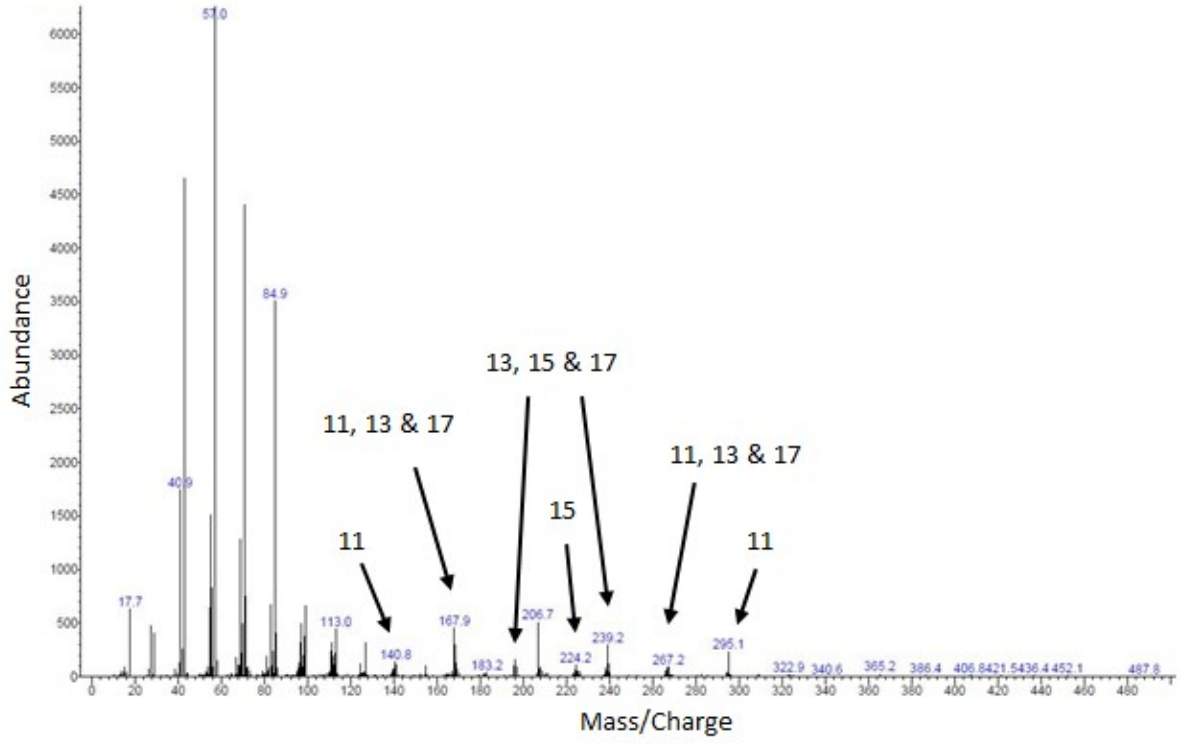

Fig. S4 Observed mass spectrum for peak 13, containing 11,15-dimethylnonacosane and 13,17-dimethylnonacosane. Numbers indicate locations of cleavage that would produce each diagnostic peak (see Table S1).

Alkenes

Most of the alkenes were identified from their mass spectrum and their corresponding molecular ion (M^+) peak. The positioning of the double bond cannot be determined from mass spectra alone because this does not produce an ion of significant stability. When possible, we confirmed double bond positions via comparison of spectra and ECL values with databases and other literature sources (Bartelt et al. 1984, 1986, Blomquist et al. 1985, Howard et al. 2001, Vaníčková et al. 2012, Curtis et al. 2013, Soares et al. 2017). For example, FID peak 2 (ECL 26.74, Fig. S5) was identified as heptacosene using these sources. The molecular ion is clear at 378 Da and the ECL is consistent with a C_{27} compound. Comparison of our observed ECL to the literature suggests that this compound contains a double bond at the 9-position

(Vaníčková et al. 2012, Curtis et al. 2013), although this cannot be definitively verified with the mass spectrum alone. Unfortunately, derivitization of these compounds was not possible due to their low absolute quantities. Peaks 8 (ECL 28.75) and 16 (ECL 30.77) was also be identified as an alkene, with the double bond also likely in the 9th position, from the mass spectrum and other literature sources. In one other case (FID peak 9, ECL 28.83), the structure of an alkene could not reasonably be inferred via ECL, as multiple isomers have similar values.

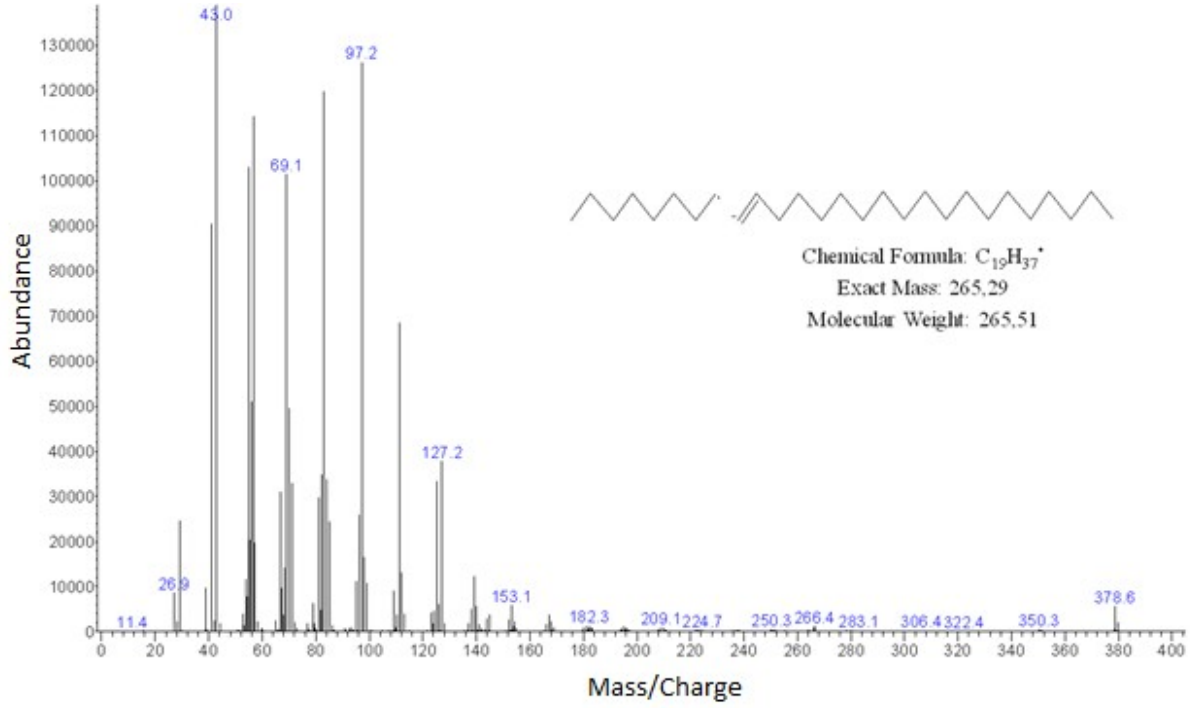

Fig S5 Observed mass spectrum of peak 2 (9-heptacosene), with stick diagram showing fragmentation pattern.

Esters

A homologous series of esters was discovered upon investigation of FID peaks 1 and 3. The mass spectrum of these two chromatographic peaks shared a base peak at 99 Da. When a single ion chromatogram (Fig. S6) was taken of this ion it revealed a family of compounds

differing in the addition of one or more CH_2 groups. Although some of these were not integrated on the FID, they were further investigated on the GC:MS. The structures of these compounds were verified by comparison to MS databases and published literature (Blomquist et al. 1972, Finidori-Logli et al. 1996, Patel et al. 2001, Howard and Baker 2003, Böröczky et al. 2008, Dweck et al. 2015, Chinta et al. 2016, Pitts-Singer et al. 2017, Merli et al. 2018).

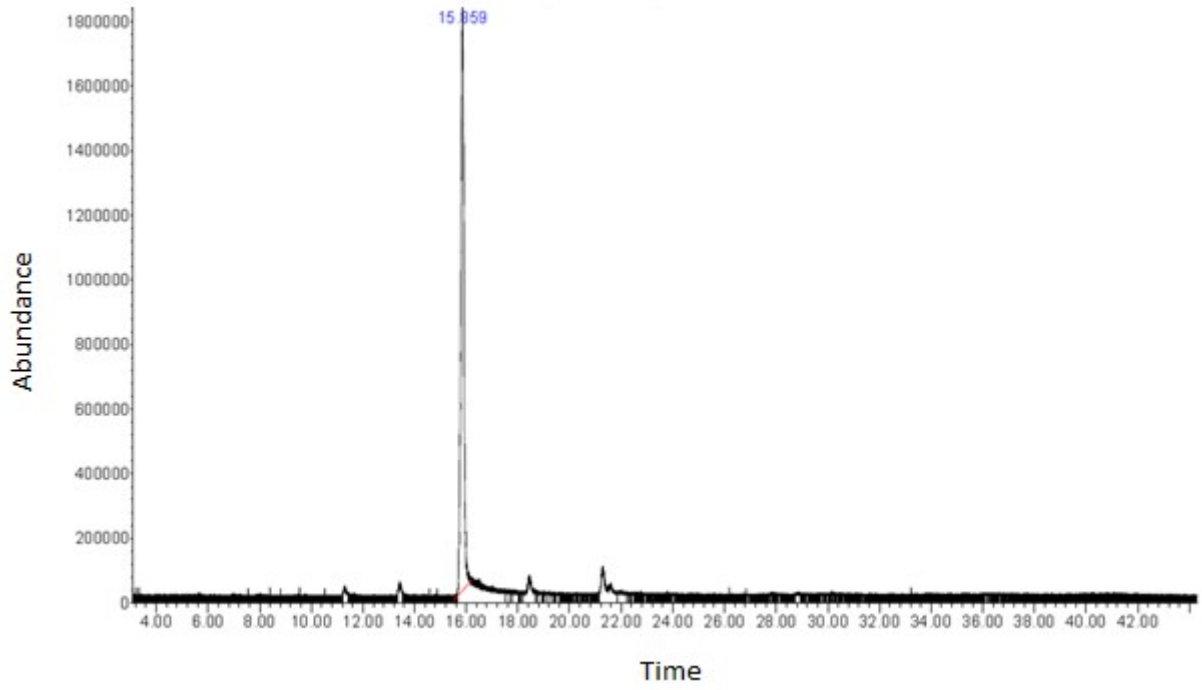

Fig. S6 Observed single ion chromatogram of mass 99 Da.

Derivatization of the sample with TMS did not alter these peaks, revealing they were neither alcohols nor carboxylic acids. The NIST MS search 2.0 GC:MS database identified these as branched esters. Using this information and a couple of diagnostic fragment ion peaks in each spectra, along with the ECL numbers, it is possible to deduce their formula, but without other sources of verification their inferred structure is not definitive. Consider FID Peak 1 as an example. Peak 1 has an ECL number of 24.97, which suggests a carbon number of 24 or 25. The M^+ ion of RCOOR′R″ esters generally becomes less prominent the longer the

R groups, and it is often not present in the mass spectrum, as in this case. The dominant peak at 99 Da seen in five of the six esters corresponds to a hexanoic acid head group produced by alpha cleavage at the carbonyl group (Fig. S7). This is characteristic of fragmentation of higher molecular weight esters. Cleavage adjacent to the alcohol oxygen produces the alkyl tail ion, which is also clearly seen in the mass spectrum of all six esters (Fig. S7). This corresponds to the other dominant ion at higher mass seen in Fig. S7 at 266 Da. It is not possible to determine exactly where the alcohol oxygen is along the alkyl chain, but the peak at 183 Da suggests it may be located at the C_6 position. This structure was also suggested by the NIST MS search database. Nevertheless, due to the lack of unequivocal evidence, we provisionally identify this compound as *x*-nonadecyl hexanoate ($C_{25}H_{50}O_2$, 382 Da).

Peaks 1a, 1c, 1d, and 3 were identified in the same way. All of these compounds produced the same 99 Da ion, representing a six-carbon ester head group, yet differed in the size of the alkyl tail ion, having more or fewer CH_2 groups. Peak 1b is an exception as it exhibited a dominant peak at 71 Da instead of 99 Da, indicating a two-carbon smaller head group. This suggests the ester has been synthesised from butyric as opposed to hexanoic acid, as illustrated in Fig. S8. Another dominant ion is present at 266 Da, indicating a C_{19} alkyl tail. Therefore, we provisionally identified this compound as *x*-nonadecyl butanoate ($C_{23}H_{46}O_2$, 354 Da).

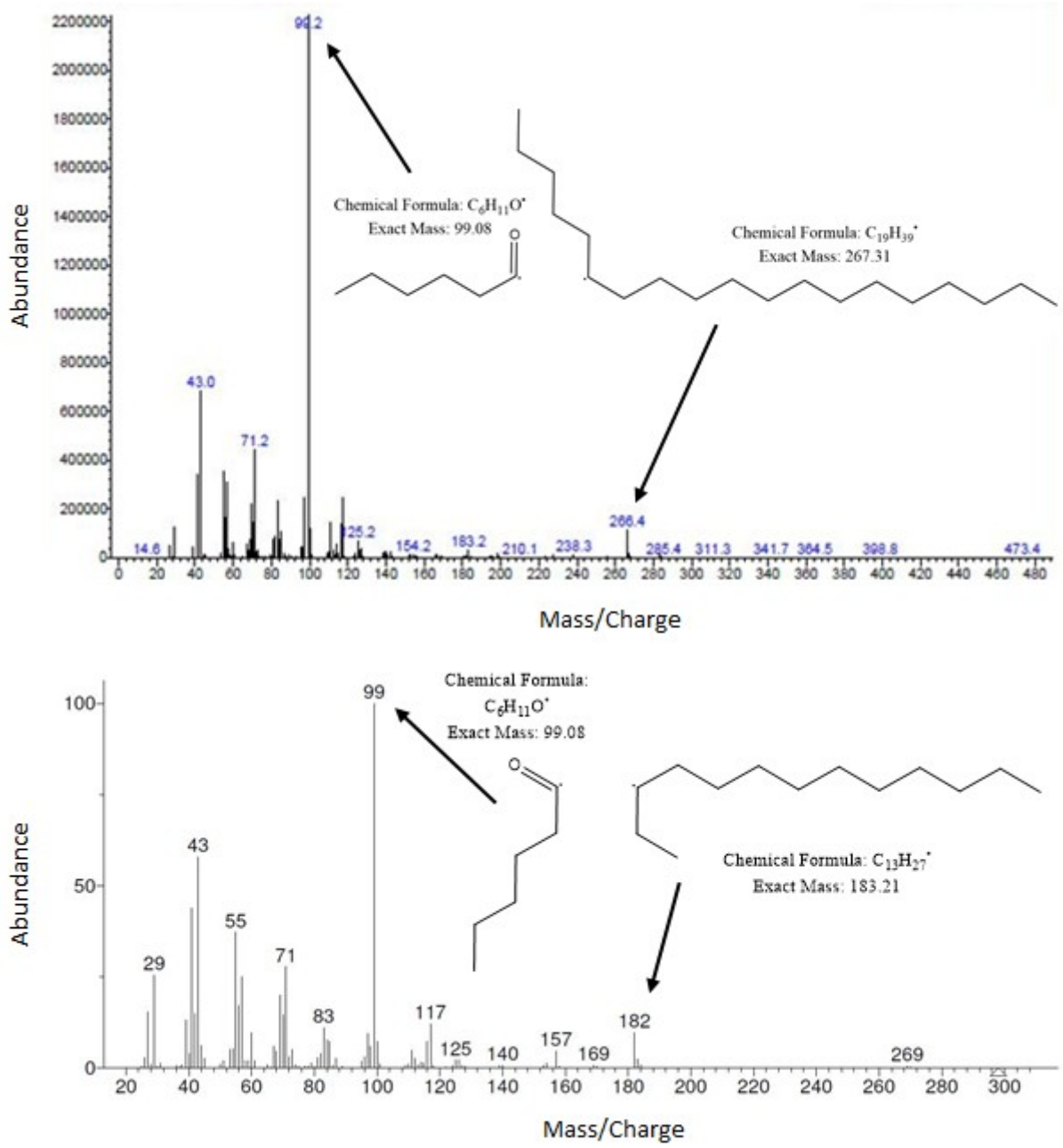

Fig. S7 Top panel, observed mass spectrum of FID peak 1 (*x*-nonadecyl hexanoate) with fragmentation pattern of proposed molecule (conjecturally drawn as 6-nonadecyl hexanoate). Bottom panel, database spectrum and fragmentation scheme of 3-tridecyl hexanoate for comparison.

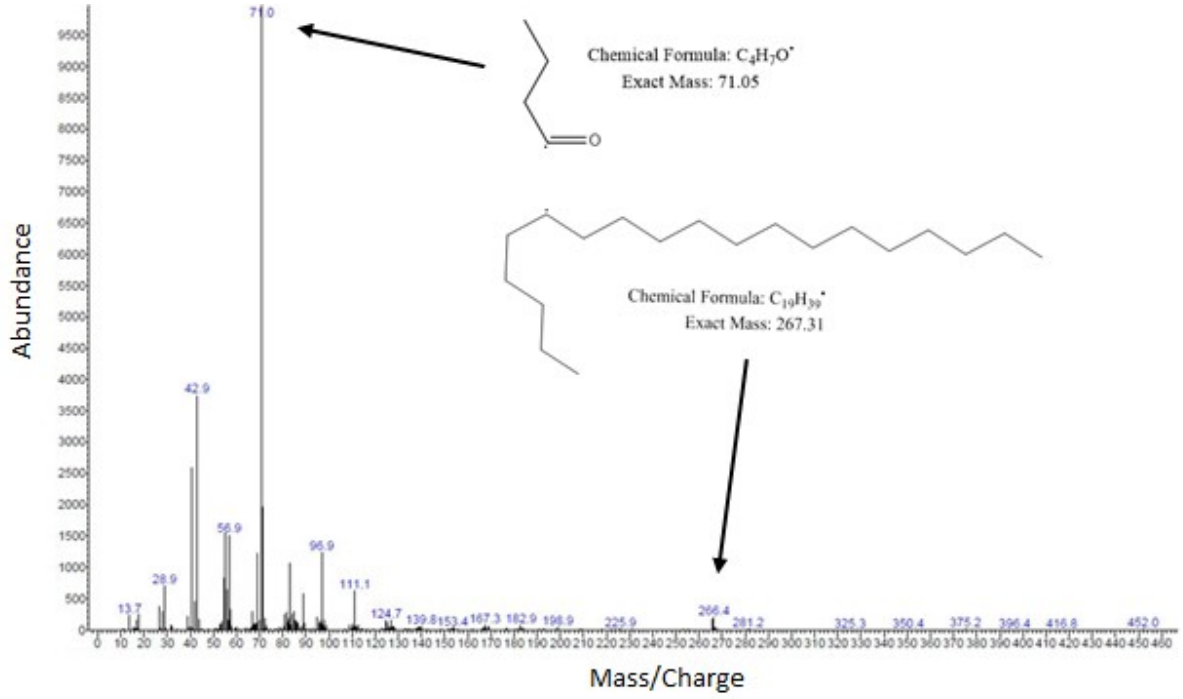

Fig. S8 Observed mass spectrum of peak 1b ester (*x*-nonadecyl butanoate), with fragmentation pattern of proposed molecule (drawn as 6-nonadecyl butanoate).

Sexual Selection Analysis

Table S2 Standardized sexual selection gradients on 17 PCs of male epicuticular compound variation across two years. Gradients and significance values for each year are based on single-year LMs and binomial GLMs, respectively, while between-year difference represents the significance of each year × PC interaction in a binomial GLM using combined data from both years. Significant effects (LRT: $P < 0.05$) are denoted in bold.

Male trait (PC)[a]	β (2013)	P (2013)	β (2017)	P (2017)	Between-year difference (P)
1	0.043	0.498	0.098	0.309	0.774
2	0.049	0.844	0.042	0.602	0.776
3	-0.130	0.126	-0.157	0.217	0.895
4	0.068	0.489	-0.035	0.898	0.494
5	**-0.419**	**< 0.001**	**-0.405**	**0.008**	0.793
6	0.209	0.084	0.098	0.787	0.690
7	0.292	0.074	**-0.835**	**0.002**	**< 0.001**
8	**0.808**	**< 0.001**	-0.219	0.339	**0.001**
9	-0.223	0.282	0.134	0.626	0.913
10	0.376	0.106	0.365	0.216	0.974
11	**-0.641**	**0.040**	**-0.620**	**0.022**	0.680
12	-0.361	0.293	-0.368	0.166	0.922
13	-0.185	0.553	**0.892**	**0.006**	**0.035**
14	-0.280	0.531	0.455	0.212	0.805
15	-0.092	0.608	**0.864**	**0.037**	0.117
16	0.062	0.952	0.027	0.965	0.777
17	-0.324	0.768	0.417	0.481	0.805

[a]See Table S4 for principal component loadings.

Table S3 Standardized intersexual selection (i.e. female mate choice) gradients on 17 PCs of male epicuticular compound variation in 2017. Estimates and significance values are based on LM and binomial GLM, respectively. Significant selection gradients (LRT: $P < 0.05$) are indicated with bold text.

Male trait (PC)[a]	β	P
1	0.125	0.277
2	0.111	0.449
3	-0.261	0.096
4	-0.113	0.372
5	**-0.474**	**0.010**
6	0.252	0.465
7	-0.486	0.166
8	0.244	0.465
9	-0.250	0.455
10	0.285	0.431
11	**-0.681**	**0.046**
12	0.379	0.359
13	**1.202**	**0.014**
14	-0.009	0.856
15	**1.590**	**0.006**
16	-0.240	0.703
17	**-2.067**	**0. 031**

[a]See Table S4 for principal component loadings

Table S4 Eigenvectors of 17 principle components of EC variation in wild male *Protopiophila litigata*. The PCA was performed on CLR-transformed relative compound abundances using the pooled 2013 and 2017 data. FID peak numbers refer to compounds identified in Table 1.

EC (FID #)	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17
% var	44.1	25.1	7.4	6.7	4.6	2.7	2.0	1.8	1.5	1.3	0.8	0.7	0.5	0.4	0.3	0.2	0.1
1	0.253	0.172	-0.028	-0.048	-0.087	-0.369	-0.544	0.021	-0.239	0.417	-0.004	-0.007	0.336	-0.054	-0.225	0.039	-0.007
2	-0.277	-0.015	0.412	0.222	0.178	-0.154	-0.204	-0.044	0.142	-0.220	-0.406	-0.174	-0.189	-0.156	-0.079	0.189	-0.174
3	0.002	-0.142	0.372	-0.125	0.180	-0.046	-0.256	-0.010	0.255	0.160	0.394	-0.321	-0.181	-0.049	0.338	-0.264	-0.321
4	0.471	-0.587	-0.082	0.527	-0.235	-0.038	0.034	0.124	0.053	-0.087	-0.046	-0.015	-0.026	0.069	0.030	-0.024	-0.015
5	0.045	-0.072	-0.154	-0.428	0.006	-0.213	-0.090	0.182	0.169	-0.094	-0.234	0.097	-0.367	0.566	-0.151	0.174	0.097
6	0.049	0.154	0.038	0.086	0.389	-0.081	0.072	0.605	-0.061	-0.283	0.277	0.350	0.072	-0.245	-0.075	0.117	0.350
7	0.004	0.285	0.172	-0.030	-0.475	-0.533	0.444	-0.143	-0.026	-0.231	0.194	-0.101	0.013	-0.035	0.044	-0.010	-0.101
8	-0.438	-0.090	0.040	0.217	0.022	-0.049	0.024	-0.158	-0.041	0.029	-0.344	0.211	0.317	0.022	-0.090	-0.192	0.211
9	-0.267	-0.378	0.009	-0.056	0.163	-0.070	0.288	-0.250	0.033	0.410	0.368	0.296	-0.019	0.054	-0.178	0.221	0.296
10	-0.061	-0.311	-0.375	-0.469	-0.015	0.030	0.034	0.016	0.192	-0.249	-0.077	-0.283	0.333	-0.405	-0.070	-0.080	-0.283
11	0.008	0.028	-0.088	-0.021	0.254	-0.004	0.164	0.019	-0.439	-0.027	-0.071	-0.068	-0.153	0.189	-0.049	-0.705	-0.068
12	-0.018	-0.017	0.043	-0.116	0.019	0.121	-0.051	-0.046	-0.244	-0.123	-0.068	0.138	0.340	0.275	0.729	0.239	0.138
13	0.334	0.118	0.243	-0.053	0.162	0.403	0.003	-0.424	-0.097	-0.314	0.152	-0.109	0.153	0.166	-0.392	0.159	-0.109
14	0.169	0.129	-0.005	-0.137	-0.253	0.204	-0.179	-0.249	0.165	-0.026	-0.096	0.603	-0.309	-0.352	0.112	-0.210	0.603
15	-0.048	0.017	0.333	-0.167	-0.363	0.445	0.236	0.435	-0.156	0.354	-0.179	-0.151	-0.027	-0.076	-0.105	0.077	-0.151
16	-0.443	0.155	-0.345	0.264	-0.329	0.257	-0.335	0.082	0.006	-0.162	0.382	-0.122	-0.099	0.174	-0.084	0.007	-0.122
17	0.064	0.137	-0.384	0.120	0.200	-0.002	0.122	-0.184	-0.318	0.173	-0.116	-0.260	-0.396	-0.303	0.168	0.346	-0.260
18	0.152	0.416	-0.201	0.213	0.184	0.099	0.240	0.024	0.606	0.273	-0.125	-0.084	0.203	0.161	0.075	-0.083	-0.084

References Cited

Bartelt, R. J., M. T. Armold, A. M. Schaner, and L. L. Jackson. **1986**. Comparative analysis of cuticular hydrocarbons in the *Drosophila virilis* species group. Comp. Biochem. Physiol. Part B Comp. Biochem. 83: 731–742.

Bartelt, R. J., T. P. Krick, and R. L. Jones. **1984**. Cuticular hydrocarbons of the yellowheaded spruce sawfly, *Pikonema alaskensis*. Insect Biochem. 14: 209–213.

Blomquist, G. J., C. L. Soliday, B. A. Byers, J. W. Brakke, and L. L. Jackson. **1972**. Cuticular lipids of insects: V. Cuticular wax esters of secondary alcohols from the grasshoppers *Melanoplus packardii* and *Melanoplus sanguinipes*. Lipids. 7: 356–362.

Blomquist, G. J., E. C. Toolson, and D. R. Nelson. **1985**. Epicuticular hydrocarbons of *Drosophila pseudoobscura* (Diptera; Drosophilidae) Identification of unusual alkadiene and alkatriene positional isomers. Insect Biochem. 15: 25–34.

Böröczky, K., K. C. Park, R. D. Minard, T. H. Jones, T. C. Baker, and J. H. Tumlinson. **2008**. Differences in cuticular lipid composition of the antennae of *Helicoverpa zea*, *Heliothis virescens*, and *Manduca sexta*. J. Insect Physiol. 54: 1385–1391.

Chinta, S. P., S. Goller, G. Uhl, and S. Schulz. **2016**. Identification and synthesis of branched wax-type esters, novel surface lipids from the spider *Argyrodes elevatus* (Araneae: Theridiidae). Chem. Biodivers. 13: 1202–1220.

Curtis, S., J. L. Sztepanacz, B. E. White, K. A. Dyer, H. D. Rundle, and P. Mayer. **2013**. Epicuticular compounds of *Drosophila subquinaria* and *D. recens*: identification, quantification, and their role in female mate choice. J. Chem. Ecol. 39: 579–590.

Dweck, H. K. M., S. A. M. Ebrahim, M. Thoma, A. A. M. Mohamed, I. W. Keesey, F. Trona, S. Lavista-Llanos, A. Svatoš, S. Sachse, M. Knaden, and B. S. Hansson. **2015**. Pheromones mediating copulation and attraction in *Drosophila*. Proc. Natl. Acad. Sci. 112: E2829–E2835.

Finidori-Logli, V., A.-G. Bagnères, D. Erdmann, W. Francke, and J.-L. Clément. **1996**. Sex recognition in *Diglyphus isaea* Walker (Hymenoptera: Eulophidae): role of an uncommon family of behaviorally active compounds. J. Chem. Ecol. 22: 2063–2079.

Howard, R. W., and J. E. Baker. **2003**. Cuticular hydrocarbons and wax esters of the ectoparasitoid *Habrobracon hebetor*: ontogenetic, reproductive, and nutritional effects. Arch. Insect Biochem. Physiol. 53: 1–18.

Howard, R. W., and G. J. Blomquist. **2005**. Ecological, behavioral, and biochemical aspects of insect hydrocarbons. Annu. Rev. Entomol. 50: 371–393.

Howard, R. W., G. Pérez-Lachaud, and J.-P. Lachaud. **2001**. Cuticular hydrocarbons of *Kapala sulcifacies* (Hymenoptera: Eucharitidae) and its host, the ponerine ant *Ectatomma ruidum* (Hymenoptera: Formicidae). Ann. Entomol. Soc. Am. 94: 707–716.

Lockey, K. H. 1988. Lipids of the insect cuticle: origin, composition and function. Comp. Biochem. Physiol. Part B Comp. Biochem. 89: 595–645.

McLafferty, F. W. 1980. Interpretation of mass spectra, 3. ed. ed, Organic chemistry series. Univ. Science Books, Mill Valley, Calif.

Merli, D., B. Mannucci, F. Bassetti, F. Corana, M. Falchetto, A. R. Malacrida, G. Gasperi, and F. Scolari. **2018**. Larval diet affects male pheromone blend in a laboratory strain of the medfly, *Ceratitis capitata* (Diptera: Tephritidae). J. Chem. Ecol. 44: 339–353.

Patel, S., D. R. Nelson, and A. G. Gibbs. **2001**. Chemical and physical analyses of wax ester properties. J. Insect Sci. 1.

Pitts-Singer, T. L., M. M. Hagen, B. R. Helm, S. Highland, J. S. Buckner, and W. P. Kemp. **2017**. Comparison of the chemical compositions of the cuticle and dufour's

gland of two solitary bee species from laboratory and field conditions. J. Chem. Ecol. 43: 451–468.

Soares, E. R. P., N. R. Batista, R. da S. Souza, V. de O. Torres, C. A. L. Cardoso, F. S. Nascimento, and W. F. Antonialli-Junior. **2017**. Variation of cuticular chemical compounds in three species of *Mischocyttarus* (Hymenoptera: Vespidae) eusocial wasps. Rev. Bras. Entomol. 61: 224–231.

Vaníčková, L., A. Svatoš, J. Kroiss, M. Kaltenpoth, R. R. do Nascimento, M. Hoskovec, R. Břízová, and B. Kalinová. **2012**. Cuticular hydrocarbons of the South American fruit fly *Anastrepha fraterculus*: variability with sex and age. J. Chem. Ecol. 38: 1133–1142.