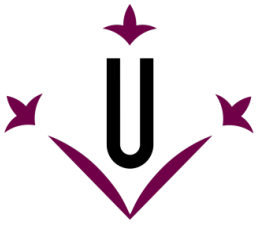


Cuticular hydrocarbons discriminate cryptic *Macrolophus* species (Hemiptera: Miridae)

César Gemeno[1*], Nerea Laserna[1], Magí Riba[2], Joan Valls[3], Cristina Castañé[4] and Oscar Alomar[4]

[1] *Department of Crop and Forest Science, University of Lleida, 25198 Lleida, Spain*

[2] *Department of Chemistry, University of Lleida, 25198 Lleida, Spain*

[3] *Biostatistics Unit. Biomedical Research Institute (IRBLLEIDA), 25198 Lleida, Spain*

[4] *IRTA Cabrils, 08348 Barcelona, Spain*

*Corresponding author

Fax: +34 973-702690

Telephone: +34 973-702531

e-mail: cesar.gemeno@pvcf.udl.cat

Macrolophus pygmaeus is commercially employed in the biological control of greenhouse and field vegetable pests. It is morphologically undistinguishable from the cryptic species *M. melanotoma*, and this interferes with the evaluation of the biological control activity of *M. pygmaeus*. We analysed the potential of cuticular hydrocarbon composition as a method to discriminate the two *Macrolophus* species. A third species, *M. costalis*, which is morphologically different from the other two species by having a dark spot at the tip of the scutellum, served as a control. Sex, diet and species, all had significant effects in the cuticular hydrocarbon profiles, but the variability associated to sex or diet was smaller than among species. Discriminant quadratic analysis of cuticular hydrocarbons confirmed the results of previous molecular genetic studies and showed, using cross-validation methods, that *M. pygmaeus* can be discriminated from *M. costalis* and *M. melanotoma* with prediction errors of 6.75% and 0%, respectively. Therefore, cuticular hydrocarbons can be used to separate *M. pygmaeus* and *M. melanotoma*, while *M. costalis* and *M. pygmaeus* can be readily distinguished from each other by the dark tip of the scutellum of the former one.

Keywords: cuticular hydrocarbons, biological control, taxonomy

Introduction

Several species in the genus *Macrolophus* (Hemiptera: Miridae) are efficient predators of vegetable crops pests (e.g. whiteflies, aphids and thrips) (Alvarado et al., 1997; Riudavets & Castañé 1998; Montserrat et al., 2000; Alomar et al., 2006; Hansen et al., 1999; Athanassiou et al., 2003; Margaritopoulos et al., 2003). Control strategies that incorporate *Macrolophus* species are based on both seasonal inoculative releases of commercially produced individuals and habitat management to conserve natural populations and enhance colonization (Alomar et al., 2002). In order to avoid the incorrect use of predator reservoir plants it is essential to correctly identify the predator species.

Several *Macrolophus* species are morphologically similar or have highly variable taxonomic characters, which hampers correct identification and questions species-identity status (Josifov, 1992; Kerzhner & Josifov, 1999). This is particularly true for the two species that have been most cited as biological control agents, *M. pygmaeus* (Rambur 1839), and *M. melanotoma* (Costa 1853) (the last one mostly cited as its junior synonym *M. caliginosus* Wagner 1951), and has probably resulted in incorrect attribution of biological control by *M. pygmaeus* to *M. melanotoma* (see Martínez-Casales et al., 2006 for a historical review). DNA methods have been employed recently to distinguish *M. pygmaeus* from *M. melanotoma* (Perdikis et al., 2003; Martínez-Casales et al., 2006). Cuticular hydrocarbon analysis could provide an additional method to separate morphologically similar species (Bagnères & Wicker-Thomas, 2010).

The insect cuticle is coated with a thin lipid layer containing a high percentage of hydrocarbons (linear, branched, saturated and unsaturated), one of which main functions is to prevent desiccation and pathogen entrance into the body (Howard & Blomquist, 2005). Cuticular hydrocarbons are usually present in high amounts and are easy to extract by quickly rinsing the specimen in non-polar organic solvents (Blomquist et al., 1987). In addition, most cuticular hydrocarbons are chemically stable and not very volatile (Martin et al., 2009). The hydrocarbon blend is often species specific and therefore it is a useful character in insect taxonomy (reviewed by Bagnères & Wicker-Thomas, 2010).

In contrast to other insect groups, there are relatively few studies of cuticular hydrocarbon taxonomy in Hemiptera. A recent publication reviews the hydrocarbons of blood-sucking bugs (Juárez & Fernández, 2007), and isolated studies compare cuticular hydrocarbons of *Orius* species (Anthocoridae) (Nakabou & Ohno, 2001), and aphids (Clements et al., 2000). With regard to mirids, analysis of cuticular hydrocarbons has been reported only for *Lygocoris pabulinus* (L.) (Drijfhout & Groot, 2001).

The purpose of the present study is to identify the cuticular hydrocarbons of three *Macrolophus* species (*M. melanotoma*, *M. pygmaeus,* and *M. costalis*), as they may be important in species recognition and mating behaviour. Since diet and sex may affect insect cuticular hydrocarbon composition (Liang & Silverman, 2000; Thomas & Simmons, 2008; Guerrieri et al., 2009), we also analysed the effect of diet (i.e., host plant and associated prey) and sex on cuticular hydrocarbons. We then determine if cuticular hydrocarbons can be used to distinguish adults of two *Macrolophus* species

(*M. melanotoma* and *M. pygmaeus*) that cannot be distinguished easily by external morphological characters alone. A third predatory species, *M. costalis* Fieber, was included in the study to confirm previous results on the taxonomic relationship of the three species (Martínez-Cascales et al., 2006).

Materials and methods

Insects

Chemical analyses were performed on adults of both sexes. Wild individuals were collected in the proximities of Cabrils (Barcelona, Spain) in the spring of 2008. *Macrolophus melanotoma* was collected from *Dittrichia viscosa* (L.) Greuter (Compositae), where it is abundant (O. A., personal observation). *Macrolophus costalis*, which is readily distinguished by the dark spot on the scutellum, was collected on *Cistus aldibus* L. (Cistaceae), where it is commonly found (O.A., personal observation). *M. pygmaeus* was obtained from a > 5-years-old laboratory colony originated from tomato fields, and maintained on tobacco plants and frozen *Ephestia kuehniella* Zeller (Lepidoptera: Pyralidae) egg prey. This colony is refreshed annually with individuals collected in local tomato fields (Cabrils, Barcelona, Spain). Examination of the length and shape of the respiratory horns of eggs (Perdikis et al., 2003; Alomar & Goula, unpub. results) was used to confirm the identity of collected females of *M. melanotoma* and *M. pygmaeus*. Some of those individuals were frozen (-20°C) immediately or within 7 days after being collected (in which case they were maintained in their respective diet), and are referred to as the “field” individuals. Additionally, in order to evaluate the effect of diet on the cuticular hydrocarbon profile, 4th-instar *M. pygmaeus* nymphs taken from the colony, and the progeny of field-collected *M. melanotoma* and *M. costalis*, were reared individually on tobacco leaves and *E. kuehniella* eggs. The resulting adults, referred to as the “laboratory” individuals, were frozen when 3 to 6 days old and analyzed at the same time as the “field” individuals described above.

Cages used for oviposition and nymphal development (7-cm diameter x 3.5-cm high with a ventilated lid) had a layer of 0.5% agar and a tobacco leaf disc placed on top of it, with the abaxial surface facing upwards, and were kept at 25 ± 2°C and 85 ± 5% RH, under a 16:8 light:dark photoregime. Individuals were moved to new cages as needed and were provided with frozen *E. kuehniella* egg prey twice a week.

Chemical Analyses

Individuals were taken out of the freezer one at a time, let defrost, and submerged in 20 µl of hexane (HPLC grade, Sigma-Aldrich, Madrid, Spain) containing 50 ng of pentadecane (98% pure, Sigma-Aldrich, Madrid, Spain) in a conical-bottom hexane-rinsed glass vial. The use of internal standard allows minimization of any differences in injection volume and in the daily response of the equipment. Previous analyses indicated that pentadecane is not present in significant amounts in cuticular extracts of these species. Five minutes after immersion in the solvent the insect was removed and the solution was either analysed immediately or returned to the freezer (-20°C) for later analysis.

For chemical analyses the volume of the extract was reduced immediately before injection to 2-4 µl with a gentle nitrogen stream and 1 µl was injected manually in a gas chromatograph (GC) in the splitless mode (split valve opened after 1 min). Samples were analysed in either a GC-FID or a GC-MSD (Agilent Technologies 6890N GC, and 5973 Network quadrupole MSD), each equipped with a DB-5 column (30 m x of 0.32 mm ID x 0.25 µm, Agilent Technologies), and run with the same temperature program:

start at 60°C for 1 min, then increase to 320°C at 10 °C/min, and maintain at this temperature for 25 min. Carrier gas was Helium at constant flow (1 ml/min), the injector was set a 250°C and the detector at 280°C.

A total of 119 insects were analysed by GC-FID, and these are the samples used in the statistical comparisons [10 individuals of each species, sex and diet (laboratory versus field), except for N=9 field *M. melanotoma* males]. Two additional insects of each sex and species were analysed by GC-MSD, and these data were used in the chemical identification of the compounds. Straight-chain alkanes were identified by comparison of retention times and mass spectra of authentic standards (Fluka, alkane standard solutions 04070 (C_8-C_{20}) and 04071 (C_{21}-C_{40}), Sigma-Aldrich, Madrid, Spain). Retention indices (RIs) were calculated according to van den Dool & Kratz (1963). Linear and methyl branched alkanes were tentatively identified by comparison of their RI and the characteristic ion fragments with those reported in the literature (Blomquist et al., 1987; Juárez & Blomquist, 1993; Carlson et al., 1998; Juárez et al., 2001; Mullen et al., 2008; Gomes et al., 2008; Dall'Aglio-Holvorcem et al., 2009).

Chain length of the selected compounds ranged between C_{24} and more than C_{40}. Since our largest alkane standard was C_{40} we fitted a curve to the straight chain alkane standards C_{27} to C_{40} to extrapolate the retention times of C_{41} and C_{42} ($R^2 = 0.99$, $y = 125.161-0.066*x+1.105*10^{-5}*x^2$, where y = estimated RT and y = chain length), and then estimated the RI of sample peaks 27 and 28 (table 1).

Statistical Analyses

Percent relative abundance [(area of the target peak/area of the internal standard peak) x 100] was the parameter used in all the group comparisons. We chose those compounds that showed a high relative abundance combined with a moderate coefficient of variation, relative to other compounds, and that occurred consistently in at least one class (i.e., species, sex, or diet). This resulted in the selection of 28 peaks, some of which were abundant in one sex or species but bellow threshold level in others. A zero value was assigned to undetected compounds. A Kolmogorov-Smirnov test was used to assess normality of the data.

Two statistical analyses were performed. The first analysis was carried out to determine the effect of diet, sex and species on the cuticular hydrocarbons. First we performed a multivariate analysis of variance (MANOVA) in order to study how the 28 cuticular hydrocarbons altogether vary with species, sex and diet. To measure the variability that could be explained by each of these factors we computed the Pillai-Bartlett statistic (Hand and Taylor, 1987), which provides a measure of the ratio of effect variance to error variance. To assess the effect of individual cuticular hydrocarbons, the three species were compared, compound by compound, males and females separately, using *ANOVA* followed by a *Tukey's* mean comparison test. In addition, the effect of sex and diet on each cuticular hydrocarbon compound was analysed with *t-tests*. *Bonferroni* method was used to correct for multiple testing so that the threshold of significance was set at 0.05/28 = 0.0017 in the *ANOVA* and *t*-tests.

The second statistical analysis was performed to build a model to predict species identity. We performed *Principal component analysis* (PCA), using the correlation matrix of untransformed data, to obtain uncorrelated scores, visualize aggregation

196 patterns of the different insect groups, and to determine which compounds had the
197 strongest effect in this pattern. To build a prediction model for the species, the main
198 principal components (those adding at least 78% total variability) were used in a
199 *Quadratic discriminant analysis* (QDA), and the prediction error of discriminant
200 functions was evaluated using a 10-fold cross-validation method. The statistical tests
201 were performed with R software (R Development Core Team, 2008). Analysis codes of
202 the prediction model are available as Supplementary Material.
203
204

205 **Results**

206

207 Representative chromatograms indicating the 28 diagnostic peaks (by order of elution)
208 are shown in fig. 1. Normality assumption was accepted for all samples ($P < 0.0017$).
209 Significant differences in cuticular hydrocarbon profiles for diet, sex and species were
210 observed (MANOVA, $P < 0.001$, table 1). The differences between sexes or diets were
211 much smaller than those between species. Using the approximated F-statistic we found
212 that variability among species is almost 4 times higher than between males and females
213 ($26.8/6.9 = 3.87$) and almost 3 times higher than between diets ($26.8/9.1 = 2.91$).

214

215 Statistical comparison of individual compounds among the three species confirmed the
216 pattern shown in fig. 1: a large proportion of compounds were similar between *M.*
217 *pygmaeus* and *M. costalis*, and different from *M. melanotoma* (fig. 2, *ANOVA*, $P <$
218 0.0017). Roughly the first half (compounds 2 to 15) were more abundant in *M.*
219 *melanotoma* than in the other two species, and the second half (compounds 16 to 28)
220 were less abundant in *M. melanotoma* than in the other species. Therefore *M.*
221 *melanotoma* is different from the other two species by having a larger quantity of the
222 early eluting compounds, and a smaller quantity of the late eluting compounds (fig. 2).
223 Sex differences in cuticular hydrocarbons were numerous in *M. melanotoma*, scarce in
224 *M. pygmaeus* and lacking in *M. costalis* (fig. 2, *t-test*, $P < 0.0017$). The cuticular
225 hydrocarbon blends of *M. melanotoma* males and females were clearly different (figs. 1
226 and 2). Individuals reared in the laboratory in isolation with a homogeneous diet of
227 tobacco and moth eggs had a higher quantity of some hydrocarbons than their “field”
228 relatives (fig. 3, *t-test*, $P < 0.0017$), but the species-specific cuticular hydrocarbon
229 pattern was not altered (fig. 3), thus indicating a weak effect by diet.

Linear alkanes, from C_{21} to C_{35}, were identified in the cuticular samples of the 3 species, but in very small quantities, and only one of them (C_{24} or tetracosane, compound 1, table 2) was abundant enough to be included among the 28 selected compounds. The rest of the hydrocarbons consisted of unbranched monoenes and alkanes with one or more methyl branches (table 2). Ion fragmentation suggests that some peaks may contain more than one compound. A characteristic compound of *M. melanotoma* is 2-methyldotriacontane (compound 7). The largest difference of *M. melanotoma* females with the other two species, besides compound 7, was a larger quantity of compounds 12 and 13 (11, 15-dimethylpentatriacontane and 11, 15, 19-trimethyl pentatriacontane). For *M. melanotoma* males the largest difference with the other two species was a larger quantity of early eluting compounds 2 to 5. *M. melanotoma* males and females differed in 25 of the 28 hydrocarbons: the first 6, by order of elution, were more abundant in males, and the other 19 were more abundant in females (fig. 2). To name some examples, males had more hentriacontene and titriacontene (compounds 3 and 6) than females, and females had more 11, 15-dimethylpentatriacontane and 11, 15, 19-trimethylheptatriacontane (compounds 12 and 20) than males. Compound 7 (2-methyldotriacontane), which is characteristic of this species, was not different between sexes.

PC1 and PC2 explained 46.28 % and 21.73 % of the variability, respectively, and clearly segregated *M. melanotoma* from the other two species, whereas sex differences were only apparent in *M. melanotoma* (fig. 4). Segregation between field and laboratory individuals was weak and only apparent in *M. melanotoma* (fig. 4), thus indicating that diet did not affect the species-specific cuticular hydrocarbon blend. Different

compounds contributed in varying degrees to the PCs (table 2). PC1, which separated *M. melanotoma* males from the other two species and from the females (fig. 4), was mainly influenced by compounds that were characteristically small in *M. melanotoma* males (compounds 17, 20, 22, 23 and 25, fig. 2). PC2, which separated *M. melanotoma* females from males and from the other two species, was strongly influenced by compounds 11 to 15 (table 2), which were more abundant in *M. melanotoma* females than in the other insects (fig. 2).

With regard to the prediction model, a QDA based on PC1, PC2 and PC3 (which together explained 78% of total variability) was used to predict the species of each test individual (see Supplementary Material files). QDA was able to correctly classify all *M. melatonoma* individuals, but 5 *M. costalis* individuals were classified as *M. pygmaeus* and 2 *M. pygmaeus* were classified as *M. costalis* (table 3). Therefore the global error was 5.88% when comparing these three species. To further assess the error prediction level, we first run a training and test method using 3 quarters of the sample (90 individuals chosen at random) to build a QDA, and the remaining quarter of the sample (29 individuals chosen at random) to test the prediction accuracy, which resulted in a 3.5% global error. Next, we used a 10-fold cross-validation method which indicated a global prediction error of 6.75% (ranging from 0% to 25%), attributed to the separation between *M. pygmaeus* and *M. costalis*, given that the error rate when discriminating *M. melatonoma* from *M. pygmaeus* or *M. costalis* was 0% in all cases. We also performed QDA separately for each sex, reaching an improved global error of 1.7% in both cases (only one insect in each sex was erroneously classified, tables S1 and S2, Supplementary Materials). Scripts of the model are available in R as Supplementary Material and can be modified to include other species and compounds.

280

Discussion

Our study shows that the cuticular hydrocarbon profiles of *M. pygmaeus* and *M. melanotoma* are markedly different from each other and can be used as a phenotypic character to discriminate adults of these otherwise practically indistinguishable species. Inclusion of a third species, *M. costalis*, which has a cuticular hydrocarbon profile that is very similar to that of *M. pygmaeus*, but different from *M. melanotoma*, strengthens the validity of the cuticular hydrocarbon method for species separation. Our results agree with sequence variation of *mt*DNA where *M. melanotoma* and *M. pygmaeus* separate in different clusters, whereas *M. pygmaeus* and *M. costalis* group in the same cluster, and are therefore considered as sister species (Martínez-Cascales et al., 2006). Crossing experiments between *M. pygmaeus* and *M. melanotoma* individuals collected from the same host plants as in this study produced no progeny (O.A. & C. C., personal observation), which is in accordance with the species status of the two taxa.

The prediction model results in 100% prediction of *M. melanotoma*, but the same model is less successful at separating *M. pygmaeus* from *M. costalis*, as the hydrocarbon profile of these two species are more similar to each other than to *M. melanotoma*. However, *M. costalis* has a dark spot on the scutellum that distinguishes it from *M. pygmaeus*, thus complete separation of all three species is possible. The same level of discrimination between *M. pygmaeus* and *M. melanotoma* than the one we have obtained is also achieved with *mt*DNA sequences (Martínez-Cascales et al., 2006), so both methods are equally suitable to distinguish between the two species. Although there were sex differences in cuticular hydrocarbon profiles, making a separate prediction model for each sex reduces only slightly the prediction error for *M. pygmaeus*

and *M. costalis*, and thus it is not necessary to analyse males and females separately. However, the two sexes are easily distinguished by eye, and so they could be analysed separately if finer discrimination is needed. Diet also had an effect on the cuticular hydrocarbons, but this effect was smaller than the differences among species, and thus diet should not affect the ability of the model to predict species identity.

To our knowledge there is only one other report of cuticular hydrocarbons in mirids (Drijhfout & Groot, 2001; Dridjhfout et al., 2003). We did not find any of the reported *L. pabulinus* compounds in *Macrolophus* cuticular extracts, neither are they present in several haematophagous triatomide heteroptera (Juárez & Fernández, 2007), however there are similarities between the cuticular hydrocarbons of triatomide bugs and *Macrolophus*. For example, the majority of compounds in both groups of insects are mostly between 20 and 40 carbons, alkenes are few, and mono and methyl branched alkanes are frequent, often with positions in carbons 11, 13, and 15. In contrast to the pattern observed in triatomide bugs and *Macrolophus*, in three species of the antocorid bug *Orius*, straight-chain alkanes, alkenes and single mono-metilated alkanes between 25 and 33 carbons in length comprise the most abundant cuticular hydrocarbons (Nakabou & Ohno, 2001).

M. melanotoma males and females had very different cuticular hydrocarbon profiles, but the hydrocarbon profiles of males and females of the other two *Macrolophus* species were very similar. Sexual differences are also relatively small in the heteropterans *Oncopeltus fasciatus* (Lygaeidae) and *Rhodnius prolixus* (Reduviidae) (Jackson, 1983; Juárez et al., 2001). Whether sexual dimorphism of cuticular hydrocarbons plays a role in mediating sexual recognition in *Macrolophus* is unknown, however there are many

examples in other insects where this is the case (reviewed in Thomas & Simmons, 2008). In the mirid *L. pabulinus* the leg hydrocarbons (Z)-9-pentacosene and (Z)-9-heptacosene are produced in sex-specific amounts and could be involved in mate location (Drijhfout & Groot, 2001; Drijhfout et al., 2003).

The effect of diet on cuticular hydrocarbon composition is well documented in social omnivorous insects such as ants and cockroaches (Liang & Silverman, 2000; Guerrieri et al., 2009). Although diet had a significant effect on the total quantity of cuticular hydrocarbons of *Macrolophus*, the species- and sex-specific patterns of cuticular hydrocarbons were not altered by the diet, suggesting that the cuticular hydrocarbon composition of *Macrolophus* is largely under genetic control and not strongly influenced by diet.

Acknowledgements

We thank Montse Lloveras for technical assistance with the chemical analysis. This research was funded in part by the Spanish Department of Science and Innovation (MICINN), projects AGL2005-03768, and AGL2006-08726 to OA and CC, and AGL2007-62366 to CG.

References

Alomar, O., Goula, M. & Albajes, R. (2002) Colonization of tomato fields by predatory mirid bugs (Hemiptera: Heteroptera) in northern Spain. *Agriculture Ecosystems and Environment* **89**, 105-115.

Alomar, O., Riudavets, J. & Castañe, C. (2006) *Macrolophus caliginosus* in the biological control of *Bemisia tabaci* on greenhouse melons. *Biological Control* **36**, 154-162.

Alvarado, P., Balta, O. & Alomar O. 1997. Efficiency of four heteroptera as predators of *Aphis gossypii* and *Macrosiphum euphorbiae* (Hom.: Aphididae). *Entomophaga* **42**, 215-226

Athanassiou C.G., Kavallieratos, N.G., Ragkou,V. S. & Buchelos C.T. (2003) Seasonal abundance and spatial distribution of the predator *Macrolophus costalis* and its prey *Myzus persicae* on tobacco. *Phytoparasitica* **31**, 8-18

Bagnères, A.-G., & Wicker-Thomas, C. (2010) Chemical taxonomy with hydrocarbons pp 121-162 in Blomquist, G. J., and Bagnères, A.-G. (Eds.) *Insect hydrocarbons: Biology, biochemistry and chemical ecology*. Cambridge, England, Cambridge University Press.

Blomquist, G. J., Nelson, D. R., & de Renobales, M. (1987) Chemistry, biochemistry, and physiology of insect cuticular lipids. *Archives of Insect Biochemistry and Physiology* **6**, 227-265

Carlson, D. A., Ulrich, R. B., & Sutton, B. D. (1998) Elution patterns from capillary GC for methyl-branched alkanes. *Journal of Chemical Ecol*ogy **24**, 1845-1865.

Clements, K.M., Wiegmann, B.M., Sorenson, C.E., Smith, C.F., Neese, P.A. & Roe, R.M. (2000) Genetic variation in the *Myzus persicae* complex (Homoptera:

Aphididae): evidence for a single species. *Annals of the Entomological Society of America* **93**, 31-46.

Dall'Aglio-Holvorcem, C. G., Benson, W. W., Gilbert L. E., Trager, J. C., & Trigo, J. R. (2009) Chemical tools to distinguish the fire ant species *Solenopsis invicta* and *S. saevissima* (Formicidae: Myrmicinae) in Southeast Brazil. *Biochemical Systematics and Ecology* **37**, 442-451.

van den Dool, H., & Kratz, P. D. (1963) A generalization of the retention index system including linear temperature programmed gas-liquid partition chromatography. *Journal of Chromatography* **11**, 463-471.

Drijfhout, F. P., & Groot, A. T. (2001) Close-range attraction in *Lygocoris pabulinus* (L.). *Journal of Chemical Ecology* **27**, 1133-1149.

Drijfhout, F. P., Groot, A. T., Beek, T. A. V., & Visser, H. (2003) Mate location in green capsid bug, *Lygocoris pabulinus*. *Entomologia Experimentalis et Applicatta* **106**, 73-77.

Gomes, C. C. G., Trigo, J. R., & Erias, A. E. (2008) Sex pheromone of the American warble fly, *Dermatobia hominis*: The role of cuticular hydrocarbons. *Journal of Chemical Ecology*, **34**, 636-646.

Guerrieri, F. J., Nehring, V., Jorgensen, C. G., Nielsen, J., Galizia, C. G., & d'Ettorre, P. (2009) Ants recognize foes, not friends. *Proceedings of the Royal Society B-Biological Sciences* **276**, 2461-2468.

Hand, D. J. and Taylor, C. C. (1987) *Multivariate Analysis of Variance and Repeated Measures*. London. Chapman and Hall.

Hansen, D. L., Brödsgaard, H. F. & Enkegaard, A. (1999) Life table characteristics of *Macrolophus caliginosus* preying upon *Tetranychus urticae*. *Entomologia Experimentalis et Applicatta* **93**, 269-275.

Howard, R. H., & Blomquist, G. J. (2005) Ecological, behavioral, and biochemical aspects of insect hydrocarbons. *Annual Review of Entomology* **50**, 371-393.

Jackson, L. L. (1983) Cuticular hydrocarbons of the milkweed bug, *Oncopeltus fasciatus* by age and sex. *Insect Biochemistry* **13**, 19-25.

Josifov, M. (1992) Zur Taxonomie der paläarktischen *Macrolophus*-Arten (Insecta, Heteroptera: Miridae). *Reichenbachia* **29**, 1-4.

Juárez, M. P., & Blomquist, G. J. (1993) Cuticular hydrocarbons of *Triatoma infestans* and *T. mazzotti*. *Comparative Biochemistry and Physiology - Part B: Biochemistry & Molecular Biology* **106**, 667-674.

Juárez, M. P., Blomquist, G. J., & Schofield, C. J. (2001) Hydrocarbons of *Rhodnius prolixus*, a Chagas disease vector. *Comparative Biochemistry and Physiology - Part B: Biochemistry & Molecular Biology* **129**, 733-746.

Juárez, M. P., & Fernández, G. C. (2007). Cuticular hydrocarbons of triatomites. *Comparative Biochemistry and Physiology - Part A: Molecular & Integrative Physiology* **147**, 711-730.

Kerzhner, I. M., & Josifov, M. (1999) Cimicomorpha II: Miridae pp. in Aukema, B. & Rieger, C. (Eds.) *Catalogue of the Heteroptera of the Palaearctic Region Vol 3*. Wageningen, The Netherlands.

Liang, D., & Silverman, J. (2000) You are what you eat: Diet modifies cuticular hydrocarbons and nestmate recognition in the Argentine ant, *Linepithema humile*. *Naturwissenschaften* **87**, 412-416.

Margaritopoulos, J. T., Tsitsipis, J. A. & Perdikis, D. C. (2003) Biological characteristics of the mirids *Macrolophus costalis* and *Macrolophus pygmaeus* preying on the tobacco form of *Myzus persicae* (Hemiptera: Aphididade). *Bulletin of Entomological Research* **93**, 39-45.

Martin, S. J., Zhong, W. H., & Drijfhout, F. P. (2009) Long-term stability of hornet cuticular hydrocarbons facilitates chemotaxonomy using museum specimens. *Biological Journal of the Linnean Society* **96**, 732-737.

Martínez-Cascales, J.I., Cenis, J.L., Cassis, G. & Sanchez, J.A. (2006) Species identity of *Macrolophus melanotoma* (Costa 1853) and *Macrolophus pygmaeus* (Rambur 1839) (Insecta: Heteroptera: Miridae) based on morphological and molecular data and bionomic implications. *Insect Systematics and Evolution* **37**, 385-404.

Montserrat, M., Albajes, R. & Castañé, C. (2000) Functional response of four Heteroptera predators preying on greenhouse whitefly and western flower thrips. *Environmental Entomology* **29**, 1075-1082.

Mullen, S. P., Millar, J. G., Schal, C., & Shaw, K. L. (2008) Identification and characterization of cuticular hydrocarbons from a rapid species radiation of hawaiian swordtailed crickets (Gryllidae : Trigonidiinae: *Laupala*). *Journal of Chemical Ecology* **34**, 198-204.

Nakabou, M., & Ohno, T. (2001) Identification of *Orius* spp. by cuticular hydrocarbons *Japanese Journal of Applied Entomology and Zoology* **45**, 18-21.

Perdikis, D. C., Margaritopoulos, J. T., Stamatis, C., Mamuris, Z., Lykouressis, D. P. Tsitsipis, J. A., & Pekas, A. (2003) Discrimination of the closely related biocontrol agents *Macrolophus melanotoma* (Hemiptera: Miridae) and *M. pygmaeus* using mitochondrial DNA analysis. *Bulletin of Entomological Research* **93**, 507-514.

R Development Core Team. (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Riudavets, J. & Castañé, C. (1998) Identification and evaluation of native predators of *Frankliniella occidentalis* (Thysanoptera, Thripidae) in the Mediterranean. *Environmental Entomol*ogy **27**, 86-93.

Thomas, M. L., & Simmons, L. W. (2008) Sexual dimorphism in cuticular hydrocarbons of the Australian field cricket *Teleogryllus oceanicus* (Orthoptera: Gryllidae). *Journal of Insect Physiology* **54**, 1081-1089.

Wheeler, A. G. (2001) (Ed.) *Biology of Plant Bugs*. Ithaca, NY, Cornell University Press.

460 Table 1. Results from MANOVA using Pillai-Bartllet's statistic when analysing
461 variability due to species, sex and diet for 28 cuticular hydrocarbons of *Macrolophus*.
462

	df	Pillai statistic	Approx. F	Numerator df	Denominator df	p-value
Species	1	1.790	26.806	56	176	0
Sex	1	0.689	6.910	28	87	1.52E-09
Diet	1	0.747	9.198	28	87	5.40E-13
Residuals	114					

463
464

465 Table 2. Chemical structure of 28 cuticular hydrocarbons of three *Macrolophus* species.
466 deduced from the characteristic ions (after GC-MSD) and the retention indices (RI)
467 relative to straight chain hydrocarbons. PC1 and PC2: percent contribution of each
468 compound to the principal components of the PCA shown in fig. 4.

469

Peak number	RT (min)	RI	Tentative chemical structure	Diagnostic ions	PC1	PC2
1	22.1	2400	*n*-tetracosane	338	0.17	0.32
2	23.34	2557	11, 15-dimethylpentacosane	168, 211, 239	0.34	0.00
3	26.82	3069	hentriacontene	434 (M+)	3.46	0.41
4	27.2	3129	11-; 13-; 15-methylhentriacontane	168, 196, 224, 252, 280, 308	4.49	0.06
5	27.39	3157	11, 19-dimethylhentriacontane	168, 196, 295, 323	5.02	0.00
6	28.02	3248	tritriacontene	462 (M+)	3.95	0.47
7	28.1	3267	2-methyldotriacontane	421 (M-43), 449 (M-15)	4.12	5.68
8	28.58	3332	11-; 13-; 15-methyltritriacontane	168, 196, 224, 252, 280, 308, 336	4.86	0.07
9	28.77	3355	mixture of di and tri-methyltritriacontanes	168, 196, 181, 211, 239, 252, 267, 323, 356	4.80	0.00
10	29.71	3465	2-methyltetratriacontane	449 (M-43), 477 (M-15)	5.42	0.03
11	29.98	3492	2, 24-dimethyltetratriacontane	168, 336, 365, 463 (M-43), 491 (M-15)	1.21	11.57
12	30.58	3552	11, 15-dimethylpentatriacontane	168, 211, 239, 280, 308, 379	0.98	12.96
13	30.81	3570	11, 15, 19-trimethylpentatriacontane	168, 239, 252, 309, 323, 393	0.58	13.20
14	31.43	3652	12, 16-; 12, 18-dimethylhexatriacontane	182, 253, 280, 308, 351, 379	0.05	9.59
15	31.68	3669	3-methylhexatriacontane	491 (M-29)	0.10	15.23
16	31.97	3692	4, 10- and 4, 12-dimethylhexatriacontane	168, 197, 225, 336, 364, 393	3.04	6.94
17	32.17	3711	4, 10, 14-trimethylhexatriacontane	168, 239, 336, 407	5.82	0.21
18	32.63	3729	11-, 13-, 15-methylheptatriacontane	168, 196, 224, 280, 308, 336, 364	1.87	8.22
19	33.03	3751	11, 15-, and 11, 19-dimethylheptatriacontane	168, 239, 280, 295, 407	2.44	8.90
20	33.35	3771	11, 15,19-trimethylhepatriacontane	168, 239, 280, 309, 351	6.73	0.17
21	33.61	3789	11, 15, 23-trimethylheptatriacontane	168, 224, 239	5.15	0.84
22	34.15	3837	unknown		6.31	0.14
23	34.48	3849	12, 16-dimethyloctatriacontane	182, 407, 253, 336, 309	6.04	1.21
24	34.84	3868	unknown	182, 225	4.85	0.65
25	36.31	3953	11, 21-dimethylnonatriacontane	168, 280, 295, 323, 407, 435	5.47	0.99
26	36.7	3973	11, 15, 19-trimethylnonatriacontane	168, 239, 309, 379, 449	5.33	0.64
27	37.89	4072	12, 16-dimethyltetracontane	182, 253, 378, 449	3.63	1.12
28	40.85	4186	11, 15, 19-trimethylhentetracontane	168, 239, 280, 308	3.78	0.38
Total					100	100

470

471

Table 3. Prediction errors in the classification of three *Macrolophus* species based on a quadratic discriminant analysis (QDA) using the three main principal components of the cuticular hydrocarbons of males and females. The number in each cell represents the number of individuals of the expected species (column) assigned to the three predicted species (rows). A total 119 insects where analysed.

	Expected species			
Predicted species by QDA	*M. melanotoma*	*M. pygmaeus*	*M. costalis*	Total
M. melanotoma	39	0	0	39
M. pygmaeus	0	38	2	40
M. costalis	0	5	35	40
Total	39	43	37	119
Prediction error (%)	0	13.16	5.4	5.88

Fig. 1. Representative gas chromatograph traces of *Macrolophus* cuticular hydrocarbons. Each trace corresponds to a different single individual. The most representative peaks were numbered 1 to 28, corresponding with the column elution time. Notice the large difference between *M. melanotoma* and the other two species as well as between *M. melanotoma* males and females. Arrow indicates the characteristic dark spot on the scutellum of *M. costalis*.

Fig. 2. Relative abundance of male and female cuticular hydrocarbons in three *Macrolophus* species. Abundance is the percentage relative to the internal standard. Asterisks above each bar pair indicate significant difference between males (grey bars) and females (empty bars) (t-test, $P < 0.0017$). Different letters for a given compound indicate significant differences among species in females (capital letters) or males (small letters) (ANOVA followed by Tukey´s test, $P < 0.0017$ in both).

Fig. 3. Effect of diet on *Macrolophus* cuticular hydrocarbons. One group ("field" individuals) fed on their natural diet (white bars), and another group ("laboratory" individuals) were reared in a standard tobacco and moth egg diet (black bars). Abundance is the percentage relative to the internal standard. Asterisks above a bar pair indicate significant difference between diets (t-test, $P < 0.0017$).

Fig. 4. Contribution of the two largest principal components (PC) from a PCA of *Macrolophus* cuticular hydrocarbons. PC1 explains 46 % of the variability and separates *M. melanotoma* males (white circles) from the females (black circles), and from the other two species (*M. pygmaeus*, triangles, *M. costalis*, squares). PC2 explains 22% of the variability and separates *M. melanotoma* females from the males and from

505 the other two species. Laboratory individuals marked with a dot, field individuals

506 unmarked.

Cuticular hydrocarbons discriminate cryptic *Macrolophus* species (Hemiptera: Miridae)

César Gemeno[1*], Nerea Laserna[1], Magí Riba[2], Joan Valls[3], Cristina Castañé[4] and Oscar Alomar[4]

[1] *Department of Crop and Forest Science, University of Lleida, 25198 Lleida, Spain*

[2] *Department of Chemistry, University of Lleida, 25198 Lleida, Spain*

[3] *Biostatistics Unit. Biomedical Research Institute (IRBLLEIDA), 25198 Lleida, Spain*

[4] *IRTA Cabrils, 08348 Barcelona, Spain*

*Corresponding author

Fax: +34 973-702690

Telephone: +34 973-702531

e-mail: cesar.gemeno@pvcf.udl.cat

Macrolophus pygmaeus is commercially employed in the biological control of greenhouse and field vegetable pests. It is morphologically undistinguishable from the cryptic species *M. melanotoma*, and this interferes with the evaluation of the biological control activity of *M. pygmaeus*. We analysed the potential of cuticular hydrocarbon composition as a method to discriminate the two *Macrolophus* species. A third species, *M. costalis*, which is morphologically different from the other two species by having a dark spot at the tip of the scutellum, served as a control. Sex, diet and species, all had significant effects in the cuticular hydrocarbon profiles, but the variability associated to sex or diet was smaller than among species. Discriminant quadratic analysis of cuticular hydrocarbons confirmed the results of previous molecular genetic studies and showed, using cross-validation methods, that *M. pygmaeus* can be discriminated from *M. costalis* and *M. melanotoma* with prediction errors of 6.75% and 0%, respectively. Therefore, cuticular hydrocarbons can be used to separate *M. pygmaeus* and *M. melanotoma*, while *M. costalis* and *M. pygmaeus* can be readily distinguished from each other by the dark tip of the scutellum of the former one.

Keywords: cuticular hydrocarbons, biological control, taxonomy

Introduction

Several species in the genus *Macrolophus* (Hemiptera: Miridae) are efficient predators of vegetable crops pests (e.g. whiteflies, aphids and thrips) (Alvarado et al., 1997; Riudavets & Castañé 1998; Montserrat et al., 2000; Alomar et al., 2006; Hansen et al., 1999; Athanassiou et al., 2003; Margaritopoulos et al., 2003). Control strategies that incorporate *Macrolophus* species are based on both seasonal inoculative releases of commercially produced individuals and habitat management to conserve natural populations and enhance colonization (Alomar et al., 2002). In order to avoid the incorrect use of predator reservoir plants it is essential to correctly identify the predator species.

Several *Macrolophus* species are morphologically similar or have highly variable taxonomic characters, which hampers correct identification and questions species-identity status (Josifov, 1992; Kerzhner & Josifov, 1999). This is particularly true for the two species that have been most cited as biological control agents, *M. pygmaeus* (Rambur 1839), and *M. melanotoma* (Costa 1853) (the last one mostly cited as its junior synonym *M. caliginosus* Wagner 1951), and has probably resulted in incorrect attribution of biological control by *M. pygmaeus* to *M. melanotoma* (see Martínez-Casales et al., 2006 for a historical review). DNA methods have been employed recently to distinguish *M. pygmaeus* from *M. melanotoma* (Perdikis et al., 2003; Martínez-Casales et al., 2006). Cuticular hydrocarbon analysis could provide an additional method to separate morphologically similar species (Bagnères & Wicker-Thomas, 2010).

The insect cuticle is coated with a thin lipid layer containing a high percentage of hydrocarbons (linear, branched, saturated and unsaturated), one of which main functions is to prevent desiccation and pathogen entrance into the body (Howard & Blomquist, 2005). Cuticular hydrocarbons are usually present in high amounts and are easy to extract by quickly rinsing the specimen in non-polar organic solvents (Blomquist et al., 1987). In addition, most cuticular hydrocarbons are chemically stable and not very volatile (Martin et al., 2009). The hydrocarbon blend is often species specific and therefore it is a useful character in insect taxonomy (reviewed by Bagnères & Wicker-Thomas, 2010).

In contrast to other insect groups, there are relatively few studies of cuticular hydrocarbon taxonomy in Hemiptera. A recent publication reviews the hydrocarbons of blood-sucking bugs (Juárez & Fernández, 2007), and isolated studies compare cuticular hydrocarbons of *Orius* species (Anthocoridae) (Nakabou & Ohno, 2001), and aphids (Clements et al., 2000). With regard to mirids, analysis of cuticular hydrocarbons has been reported only for *Lygocoris pabulinus* (L.) (Drijfhout & Groot, 2001).

The purpose of the present study is to identify the cuticular hydrocarbons of three *Macrolophus* species (*M. melanotoma*, *M. pygmaeus*, and *M. costalis*), as they may be important in species recognition and mating behaviour. Since diet and sex may affect insect cuticular hydrocarbon composition (Liang & Silverman, 2000; Thomas & Simmons, 2008; Guerrieri et al., 2009), we also analysed the effect of diet (i.e., host plant and associated prey) and sex on cuticular hydrocarbons. We then determine if cuticular hydrocarbons can be used to distinguish adults of two *Macrolophus* species

(*M. melanotoma* and *M. pygmaeus*) that cannot be distinguished easily by external morphological characters alone. A third predatory species, *M. costalis* Fieber, was included in the study to confirm previous results on the taxonomic relationship of the three species (Martínez-Cascales et al., 2006).

Materials and methods

Insects

Chemical analyses were performed on adults of both sexes. Wild individuals were collected in the proximities of Cabrils (Barcelona, Spain) in the spring of 2008. *Macrolophus melanotoma* was collected from *Dittrichia viscosa* (L.) Greuter (Compositae), where it is abundant (O. A., personal observation). *Macrolophus costalis*, which is readily distinguished by the dark spot on the scutellum, was collected on *Cistus aldibus* L. (Cistaceae), where it is commonly found (O.A., personal observation). *M. pygmaeus* was obtained from a > 5-years-old laboratory colony originated from tomato fields, and maintained on tobacco plants and frozen *Ephestia kuehniella* Zeller (Lepidoptera: Pyralidae) egg prey. This colony is refreshed annually with individuals collected in local tomato fields (Cabrils, Barcelona, Spain). Examination of the length and shape of the respiratory horns of eggs (Perdikis et al., 2003; Alomar & Goula, unpub. results) was used to confirm the identity of collected females of *M. melanotoma* and *M. pygmaeus*. Some of those individuals were frozen (-20°C) immediately or within 7 days after being collected (in which case they were maintained in their respective diet), and are referred to as the "field" individuals. Additionally, in order to evaluate the effect of diet on the cuticular hydrocarbon profile, 4th-instar *M. pygmaeus* nymphs taken from the colony, and the progeny of field-collected *M. melanotoma* and *M. costalis*, were reared individually on tobacco leaves and *E. kuehniella* eggs. The resulting adults, referred to as the "laboratory" individuals, were frozen when 3 to 6 days old and analyzed at the same time as the "field" individuals described above.

Cages used for oviposition and nymphal development (7-cm diameter x 3.5-cm high with a ventilated lid) had a layer of 0.5% agar and a tobacco leaf disc placed on top of it, with the abaxial surface facing upwards, and were kept at 25 ± 2°C and 85 ± 5% RH, under a 16:8 light:dark photoregime. Individuals were moved to new cages as needed and were provided with frozen *E. kuehniella* egg prey twice a week.

Chemical Analyses

Individuals were taken out of the freezer one at a time, let defrost, and submerged in 20 µl of hexane (HPLC grade, Sigma-Aldrich, Madrid, Spain) containing 50 ng of pentadecane (98% pure, Sigma-Aldrich, Madrid, Spain) in a conical-bottom hexane-rinsed glass vial. The use of internal standard allows minimization of any differences in injection volume and in the daily response of the equipment. Previous analyses indicated that pentadecane is not present in significant amounts in cuticular extracts of these species. Five minutes after immersion in the solvent the insect was removed and the solution was either analysed immediately or returned to the freezer (-20°C) for later analysis.

For chemical analyses the volume of the extract was reduced immediately before injection to 2-4 µl with a gentle nitrogen stream and 1 µl was injected manually in a gas chromatograph (GC) in the splitless mode (split valve opened after 1 min). Samples were analysed in either a GC-FID or a GC-MSD (Agilent Technologies 6890N GC, and 5973 Network quadrupole MSD), each equipped with a DB-5 column (30 m x of 0.32 mm ID x 0.25 µm, Agilent Technologies), and run with the same temperature program:

start at 60°C for 1 min, then increase to 320°C at 10 °C/min, and maintain at this temperature for 25 min. Carrier gas was Helium at constant flow (1 ml/min), the injector was set a 250°C and the detector at 280°C.

A total of 119 insects were analysed by GC-FID, and these are the samples used in the statistical comparisons [10 individuals of each species, sex and diet (laboratory versus field), except for N=9 field *M. melanotoma* males]. Two additional insects of each sex and species were analysed by GC-MSD, and these data were used in the chemical identification of the compounds. Straight-chain alkanes were identified by comparison of retention times and mass spectra of authentic standards (Fluka, alkane standard solutions 04070 (C_8-C_{20}) and 04071 (C_{21}-C_{40}), Sigma-Aldrich, Madrid, Spain). Retention indices (RIs) were calculated according to van den Dool & Kratz (1963). Linear and methyl branched alkanes were tentatively identified by comparison of their RI and the characteristic ion fragments with those reported in the literature (Blomquist et al., 1987; Juárez & Blomquist, 1993; Carlson et al., 1998; Juárez et al., 2001; Mullen et al., 2008; Gomes et al., 2008; Dall'Aglio-Holvorcem et al., 2009).

Chain length of the selected compounds ranged between C_{24} and more than C_{40}. Since our largest alkane standard was C_{40} we fitted a curve to the straight chain alkane standards C_{27} to C_{40} to extrapolate the retention times of C_{41} and C_{42} ($R^2 = 0.99$, $y = 125.161-0.066*x+1.105*10^{-5}*x^2$, where y = estimated RT and y = chain length), and then estimated the RI of sample peaks 27 and 28 (table 1).

Statistical Analyses

Percent relative abundance [(area of the target peak/area of the internal standard peak) x 100] was the parameter used in all the group comparisons. We chose those compounds that showed a high relative abundance combined with a moderate coefficient of variation, relative to other compounds, and that occurred consistently in at least one class (i.e., species, sex, or diet). This resulted in the selection of 28 peaks, some of which were abundant in one sex or species but bellow threshold level in others. A zero value was assigned to undetected compounds. A Kolmogorov-Smirnov test was used to assess normality of the data.

Two statistical analyses were performed. The first analysis was carried out to determine the effect of diet, sex and species on the cuticular hydrocarbons. First we performed a multivariate analysis of variance (MANOVA) in order to study how the 28 cuticular hydrocarbons altogether vary with species, sex and diet. To measure the variability that could be explained by each of these factors we computed the Pillai-Bartlett statistic (Hand and Taylor, 1987), which provides a measure of the ratio of effect variance to error variance. To assess the effect of individual cuticular hydrocarbons, the three species were compared, compound by compound, males and females separately, using *ANOVA* followed by a *Tukey's* mean comparison test. In addition, the effect of sex and diet on each cuticular hydrocarbon compound was analysed with *t-tests*. *Bonferroni* method was used to correct for multiple testing so that the threshold of significance was set at 0.05/28 = 0.0017 in the *ANOVA* and *t*-tests.

The second statistical analysis was performed to build a model to predict species identity. We performed *Principal component analysis* (PCA), using the correlation matrix of untransformed data, to obtain uncorrelated scores, visualize aggregation

196 patterns of the different insect groups, and to determine which compounds had the
197 strongest effect in this pattern. To build a prediction model for the species, the main
198 principal components (those adding at least 78% total variability) were used in a
199 *Quadratic discriminant analysis* (QDA), and the prediction error of discriminant
200 functions was evaluated using a 10-fold cross-validation method. The statistical tests
201 were performed with R software (R Development Core Team, 2008). Analysis codes of
202 the prediction model are available as Supplementary Material.
203
204

205 **Results**

206

207 Representative chromatograms indicating the 28 diagnostic peaks (by order of elution)
208 are shown in fig. 1. Normality assumption was accepted for all samples ($P < 0.0017$).
209 Significant differences in cuticular hydrocarbon profiles for diet, sex and species were
210 observed (MANOVA, $P < 0.001$, table 1). The differences between sexes or diets were
211 much smaller than those between species. Using the approximated F-statistic we found
212 that variability among species is almost 4 times higher than between males and females
213 (26.8/6.9 = 3.87) and almost 3 times higher than between diets (26.8/9.1= 2.91).

214

215 Statistical comparison of individual compounds among the three species confirmed the
216 pattern shown in fig. 1: a large proportion of compounds were similar between *M.*
217 *pygmaeus* and *M. costalis*, and different from *M. melanotoma* (fig. 2, *ANOVA*, $P <$
218 0.0017). Roughly the first half (compounds 2 to 15) were more abundant in *M.*
219 *melanotoma* than in the other two species, and the second half (compounds 16 to 28)
220 were less abundant in *M. melanotoma* than in the other species. Therefore *M.*
221 *melanotoma* is different from the other two species by having a larger quantity of the
222 early eluting compounds, and a smaller quantity of the late eluting compounds (fig. 2).
223 Sex differences in cuticular hydrocarbons were numerous in *M. melanotoma*, scarce in
224 *M. pygmaeus* and lacking in *M. costalis* (fig. 2, *t-test*, $P < 0.0017$). The cuticular
225 hydrocarbon blends of *M. melanotoma* males and females were clearly different (figs. 1
226 and 2). Individuals reared in the laboratory in isolation with a homogeneous diet of
227 tobacco and moth eggs had a higher quantity of some hydrocarbons than their "field"
228 relatives (fig. 3, *t-test*, $P < 0.0017$), but the species-specific cuticular hydrocarbon
229 pattern was not altered (fig. 3), thus indicating a weak effect by diet.

Linear alkanes, from C_{21} to C_{35}, were identified in the cuticular samples of the 3 species, but in very small quantities, and only one of them (C_{24} or tetracosane, compound 1, table 2) was abundant enough to be included among the 28 selected compounds. The rest of the hydrocarbons consisted of unbranched monoenes and alkanes with one or more methyl branches (table 2). Ion fragmentation suggests that some peaks may contain more than one compound. A characteristic compound of *M. melanotoma* is 2-methyldotriacontane (compound 7). The largest difference of *M. melanotoma* females with the other two species, besides compound 7, was a larger quantity of compounds 12 and 13 (11, 15-dimethylpentatriacontane and 11, 15, 19-trimethyl pentatriacontane). For *M. melanotoma* males the largest difference with the other two species was a larger quantity of early eluting compounds 2 to 5. *M. melanotoma* males and females differed in 25 of the 28 hydrocarbons: the first 6, by order of elution, were more abundant in males, and the other 19 were more abundant in females (fig. 2). To name some examples, males had more hentriacontene and titriacontene (compounds 3 and 6) than females, and females had more 11, 15-dimethylpentatriacontane and 11, 15, 19-trimethylheptatriacontane (compounds 12 and 20) than males. Compound 7 (2-methyldotriacontane), which is characteristic of this species, was not different between sexes.

PC1 and PC2 explained 46.28 % and 21.73 % of the variability, respectively, and clearly segregated *M. melanotoma* from the other two species, whereas sex differences were only apparent in *M. melanotoma* (fig. 4). Segregation between field and laboratory individuals was weak and only apparent in *M. melanotoma* (fig. 4), thus indicating that diet did not affect the species-specific cuticular hydrocarbon blend. Different

compounds contributed in varying degrees to the PCs (table 2). PC1, which separated *M. melanotoma* males from the other two species and from the females (fig. 4), was mainly influenced by compounds that were characteristically small in *M. melanotoma* males (compounds 17, 20, 22, 23 and 25, table 2, fig. 2). PC2, which separated *M. melanotoma* females from males and from the other two species, was strongly influenced by compounds 11 to 15 (table 2), which were more abundant in *M. melanotoma* females than in the other insects (fig. 2).

With regard to the prediction model, a QDA based on PC1, PC2 and PC3 (which together explained 78% of total variability) was used to predict the species of each test individual (see Supplementary Material files). QDA was able to correctly classify all *M. melatonoma* individuals, but 5 *M. costalis* individuals were classified as *M. pygmaeus* and 2 *M. pygmaeus* were classified as *M. costalis* (table 3). Therefore the global error was 5.88% when comparing these three species. To further assess the error prediction level, we first run a training and test method using 3 quarters of the sample (90 individuals chosen at random) to build a QDA, and the remaining quarter of the sample (29 individuals chosen at random) to test the prediction accuracy, which resulted in a 3.5% global error. Next, we used a 10-fold cross-validation method which indicated a global prediction error of 6.75% (ranging from 0% to 25%), attributed to the separation between *M. pygmaeus* and *M. costalis*, given that the error rate when discriminating *M. melatonoma* from *M. pygmaeus* or *M. costalis* was 0% in all cases. We also performed QDA separately for each sex, reaching an improved global error of 1.7% in both cases (only one insect in each sex was erroneously classified, tables S1 and S2, Supplementary Materials). Scripts of the model are available in R as Supplementary Material and can be modified to include other species and compounds.

280

Discussion

Our study shows that the cuticular hydrocarbon profiles of *M. pygmaeus* and *M. melanotoma* are markedly different from each other and can be used as a phenotypic character to discriminate adults of these otherwise practically indistinguishable species. Inclusion of a third species, *M. costalis*, which has a cuticular hydrocarbon profile that is very similar to that of *M. pygmaeus*, but different from *M. melanotoma*, strengthens the validity of the cuticular hydrocarbon method for species separation. Our results agree with sequence variation of *mt*DNA where *M. melanotoma* and *M. pygmaeus* separate in different clusters, whereas *M. pygmaeus* and *M. costalis* group in the same cluster, and are therefore considered as sister species (Martínez-Cascales et al., 2006). Crossing experiments between *M. pygmaeus* and *M. melanotoma* individuals collected from the same host plants as in this study produced no progeny (O.A. & C. C., personal observation), which is in accordance with the species status of the two taxa.

The prediction model results in 100% prediction of *M. melanotoma*, but the same model is less successful at separating *M. pygmaeus* from *M. costalis*, as the hydrocarbon profile of these two species are more similar to each other than to *M. melanotoma*. However, *M. costalis* has a dark spot on the scutellum that distinguishes it from *M. pygmaeus*, thus complete separation of all three species is possible. The same level of discrimination between *M. pygmaeus* and *M. melanotoma* than the one we have obtained is also achieved with *mt*DNA sequences (Martínez-Cascales et al., 2006), so both methods are equally suitable to distinguish between the two species. Although there were sex differences in cuticular hydrocarbon profiles, making a separate prediction model for each sex reduces only slightly the prediction error for *M. pygmaeus*

306 and *M. costalis*, and thus it is not necessary to analyse males and females separately.
307 However, the two sexes are easily distinguished by eye, and so they could be analysed
308 separately if finer discrimination is needed. Diet also had an effect on the cuticular
309 hydrocarbons, but this effect was smaller than the differences among species, and thus
310 diet should not affect the ability of the model to predict species identity.
311
312 To our knowledge there is only one other report of cuticular hydrocarbons in mirids
313 (Drijhfout & Groot, 2001; Dridjhfout et al., 2003). We did not find any of the reported
314 *L. pabulinus* compounds in *Macrolophus* cuticular extracts, neither are they present in
315 several haematophagous triatomide heteroptera (Juárez & Fernández, 2007), however
316 there are similarities between the cuticular hydrocarbons of triatomide bugs and
317 *Macrolophus*. For example, the majority of compounds in both groups of insects are
318 mostly between 20 and 40 carbons, alkenes are few, and mono and methyl branched
319 alkanes are frequent, often with positions in carbons 11, 13, and 15. In contrast to the
320 pattern observed in triatomide bugs and *Macrolophus*, in three species of the antocorid
321 bug *Orius*, straight-chain alkanes, alkenes and single mono-metilated alkanes between
322 25 and 33 carbons in length comprise the most abundant cuticular hydrocarbons
323 (Nakabou & Ohno, 2001).
324
325 *M. melanotoma* males and females had very different cuticular hydrocarbon profiles,
326 but the hydrocarbon profiles of males and females of the other two *Macrolophus* species
327 were very similar. Sexual differences are also relatively small in the heteropterans
328 *Oncopeltus fasciatus* (Lygaeidae) and *Rhodnius prolixus* (Reduviidae) (Jackson, 1983;
329 Juárez et al., 2001). Whether sexual dimorphism of cuticular hydrocarbons plays a role
330 in mediating sexual recognition in *Macrolophus* is unknown, however there are many

examples in other insects where this is the case (reviewed in Thomas & Simmons, 2008). In the mirid *L. pabulinus* the leg hydrocarbons (Z)-9-pentacosene and (Z)-9-heptacosene are produced in sex-specific amounts and could be involved in mate location (Drijhfout & Groot, 2001; Drijhfout et al., 2003).

The effect of diet on cuticular hydrocarbon composition is well documented in social omnivorous insects such as ants and cockroaches (Liang & Silverman, 2000; Guerrieri et al., 2009). Although diet had a significant effect on the total quantity of cuticular hydrocarbons of *Macrolophus*, the species- and sex-specific patterns of cuticular hydrocarbons were not altered by the diet, suggesting that the cuticular hydrocarbon composition of *Macrolophus* is largely under genetic control and not strongly influenced by diet.

Acknowledgements

We thank Montse Lloveras for technical assistance with the chemical analysis. This research was funded in part by the Spanish Department of Science and Innovation (MICINN), projects AGL2005-03768, and AGL2006-08726 to OA and CC, and AGL2007-62366 to CG.

References

Alomar, O., Goula, M. & Albajes, R. (2002) Colonization of tomato fields by predatory mirid bugs (Hemiptera: Heteroptera) in northern Spain. *Agriculture Ecosystems and Environment* **89**, 105-115.

Alomar, O., Riudavets, J. & Castañe, C. (2006) *Macrolophus caliginosus* in the biological control of *Bemisia tabaci* on greenhouse melons. *Biological Control* **36**, 154-162.

Alvarado, P., Balta, O. & Alomar O. 1997. Efficiency of four heteroptera as predators of *Aphis gossypii* and *Macrosiphum euphorbiae* (Hom.: Aphididae). *Entomophaga* **42**, 215-226

Athanassiou C.G., Kavallieratos, N.G., Ragkou,V. S. & Buchelos C.T. (2003) Seasonal abundance and spatial distribution of the predator *Macrolophus costalis* and its prey *Myzus persicae* on tobacco. *Phytoparasitica* **31**, 8-18

Bagnères, A.-G., & Wicker-Thomas, C. (2010) Chemical taxonomy with hydrocarbons pp 121-162 in Blomquist, G. J., and Bagnères, A.-G. (Eds.) *Insect hydrocarbons: Biology, biochemistry and chemical ecology*. Cambridge, England, Cambridge University Press.

Blomquist, G. J., Nelson, D. R., & de Renobales, M. (1987) Chemistry, biochemistry, and physiology of insect cuticular lipids. *Archives of Insect Biochemistry and Physiology* **6**, 227-265

Carlson, D. A., Ulrich, R. B., & Sutton, B. D. (1998) Elution patterns from capillary GC for methyl-branched alkanes. *Journal of Chemical Ecol*ogy **24**, 1845-1865.

Clements, K.M., Wiegmann, B.M., Sorenson, C.E., Smith, C.F., Neese, P.A. & Roe, R.M. (2000) Genetic variation in the *Myzus persicae* complex (Homoptera:

Aphididae): evidence for a single species. *Annals of the Entomological Society of America* **93**, 31-46.

Dall'Aglio-Holvorcem, C. G., Benson, W. W., Gilbert L. E., Trager, J. C., & Trigo, J. R. (2009) Chemical tools to distinguish the fire ant species *Solenopsis invicta* and *S. saevissima* (Formicidae: Myrmicinae) in Southeast Brazil. *Biochemical Systematics and Ecology* **37**, 442-451.

van den Dool, H., & Kratz, P. D. (1963) A generalization of the retention index system including linear temperature programmed gas-liquid partition chromatography. *Journal of Chromatography* **11**, 463-471.

Drijfhout, F. P., & Groot, A. T. (2001) Close-range attraction in *Lygocoris pabulinus* (L.). *Journal of Chemical Ecology* **27**, 1133-1149.

Drijfhout, F. P., Groot, A. T., Beek, T. A. V., & Visser, H. (2003) Mate location in green capsid bug, *Lygocoris pabulinus*. *Entomologia Experimentalis et Applicatta* **106**, 73-77.

Gomes, C. C. G., Trigo, J. R., & Erias, A. E. (2008) Sex pheromone of the American warble fly, *Dermatobia hominis*: The role of cuticular hydrocarbons. *Journal of Chemical Ecology*, **34**, 636-646.

Guerrieri, F. J., Nehring, V., Jorgensen, C. G., Nielsen, J., Galizia, C. G., & d'Ettorre, P. (2009) Ants recognize foes, not friends. *Proceedings of the Royal Society B-Biological Sciences* **276**, 2461-2468.

Hand, D. J. and Taylor, C. C. (1987) *Multivariate Analysis of Variance and Repeated Measures*. London. Chapman and Hall.

Hansen, D. L., Brödsgaard, H. F. & Enkegaard, A. (1999) Life table characteristics of *Macrolophus caliginosus* preying upon *Tetranychus urticae*. *Entomologia Experimentalis et Applicatta* **93**, 269-275.

Howard, R. H., & Blomquist, G. J. (2005) Ecological, behavioral, and biochemical aspects of insect hydrocarbons. *Annual Review of Entomology* **50**, 371-393.

Jackson, L. L. (1983) Cuticular hydrocarbons of the milkweed bug, *Oncopeltus fasciatus* by age and sex. *Insect Biochemistry* **13**, 19-25.

Josifov, M. (1992) Zur Taxonomie der paläarktischen *Macrolophus*-Arten (Insecta, Heteroptera: Miridae). *Reichenbachia* **29**, 1-4.

Juárez, M. P., & Blomquist, G. J. (1993) Cuticular hydrocarbons of *Triatoma infestans* and *T. mazzotti*. *Comparative Biochemistry and Physiology - Part B: Biochemistry & Molecular Biology* **106**, 667-674.

Juárez, M. P., Blomquist, G. J., & Schofield, C. J. (2001) Hydrocarbons of *Rhodnius prolixus*, a Chagas disease vector. *Comparative Biochemistry and Physiology - Part B: Biochemistry & Molecular Biology* **129**, 733-746.

Juárez, M. P., & Fernández, G. C. (2007). Cuticular hydrocarbons of triatomites. *Comparative Biochemistry and Physiology - Part A: Molecular & Integrative Physiology* **147**, 711-730.

Kerzhner, I. M., & Josifov, M. (1999) Cimicomorpha II: Miridae pp. in Aukema, B. & Rieger, C. (Eds.) *Catalogue of the Heteroptera of the Palaearctic Region Vol 3*. Wageningen, The Netherlands.

Liang, D., & Silverman, J. (2000) You are what you eat: Diet modifies cuticular hydrocarbons and nestmate recognition in the Argentine ant, *Linepithema humile*. *Naturwissenschaften* **87**, 412-416.

Margaritopoulos, J. T., Tsitsipis, J. A. & Perdikis, D. C. (2003) Biological characteristics of the mirids *Macrolophus costalis* and *Macrolophus pygmaeus* preying on the tobacco form of *Myzus persicae* (Hemiptera: Aphididade). *Bulletin of Entomological Research* **93**, 39-45.

Martin, S. J., Zhong, W. H., & Drijfhout, F. P. (2009) Long-term stability of hornet cuticular hydrocarbons facilitates chemotaxonomy using museum specimens. *Biological Journal of the Linnean Society* **96**, 732-737.

Martínez-Cascales, J.I., Cenis, J.L., Cassis, G. & Sanchez, J.A. (2006) Species identity of *Macrolophus melanotoma* (Costa 1853) and *Macrolophus pygmaeus* (Rambur 1839) (Insecta: Heteroptera: Miridae) based on morphological and molecular data and bionomic implications. *Insect Systematics and Evolution* **37**, 385-404.

Montserrat, M., Albajes, R. & Castañé, C. (2000) Functional response of four Heteroptera predators preying on greenhouse whitefly and western flower thrips. *Environmental Entomology* **29**, 1075-1082.

Mullen, S. P., Millar, J. G., Schal, C., & Shaw, K. L. (2008) Identification and characterization of cuticular hydrocarbons from a rapid species radiation of hawaiian swordtailed crickets (Gryllidae : Trigonidiinae: *Laupala*). *Journal of Chemical Ecology* **34**, 198-204.

Nakabou, M., & Ohno, T. (2001) Identification of *Orius* spp. by cuticular hydrocarbons *Japanese Journal of Applied Entomology and Zoology* **45**, 18-21.

Perdikis, D. C., Margaritopoulos, J. T., Stamatis, C., Mamuris, Z., Lykouressis, D. P. Tsitsipis, J. A., & Pekas, A. (2003) Discrimination of the closely related biocontrol agents *Macrolophus melanotoma* (Hemiptera: Miridae) and *M. pygmaeus* using mitochondrial DNA analysis. *Bulletin of Entomological Research* **93**, 507-514.

R Development Core Team. (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Riudavets, J. & Castañé, C. (1998) Identification and evaluation of native predators of *Frankliniella occidentalis* (Thysanoptera, Thripidae) in the Mediterranean. *Environmental Entomol*ogy **27**, 86-93.

Thomas, M. L., & Simmons, L. W. (2008) Sexual dimorphism in cuticular hydrocarbons of the Australian field cricket *Teleogryllus oceanicus* (Orthoptera: Gryllidae). *Journal of Insect Physiology* **54**, 1081-1089.

Wheeler, A. G. (2001) (Ed.) *Biology of Plant Bugs*. Ithaca, NY, Cornell University Press.

460 Table 1. Results from MANOVA using Pillai-Bartllet's statistic when analysing

461 variability due to species, sex and diet for 28 cuticular hydrocarbons of *Macrolophus*.

462

	df	Pillai statistic	Approx. F	Numerator df	Denominator df	p-value
Species	1	1.790	26.806	56	176	0
Sex	1	0.689	6.910	28	87	1.52E-09
Diet	1	0.747	9.198	28	87	5.40E-13
Residuals	114					

463

464

Table 2. Chemical structure of 28 cuticular hydrocarbons of three *Macrolophus* species. deduced from the characteristic ions (after GC-MSD) and the retention indices (RI) relative to straight chain hydrocarbons. PC1 and PC2: percent contribution of each compound to the principal components of the PCA shown in fig. 4.

Peak number	RT (min)	RI	Tentative chemical structure	Diagnostic ions	PC1	PC2
1	22.1	2400	*n*-tetracosane	338	0.17	0.32
2	23.34	2557	11, 15-dimethylpentacosane	168, 211, 239	0.34	0.00
3	26.82	3069	hentriacontene	434 (M+)	3.46	0.41
4	27.2	3129	11-; 13-; 15-methylhentriacontane	168, 196, 224, 252, 280, 308	4.49	0.06
5	27.39	3157	11, 19-dimethylhentriacontane	168, 196, 295, 323	5.02	0.00
6	28.02	3248	tritriacontene	462 (M+)	3.95	0.47
7	28.1	3267	2-methyldotriacontane	421 (M-43), 449 (M-15)	4.12	5.68
8	28.58	3332	11-; 13-; 15-methyltritriacontane	168, 196, 224, 252, 280, 308, 336	4.86	0.07
9	28.77	3355	mixture of di and tri-methyltritriacontanes	168, 196, 181, 211, 239, 252, 267, 323, 356	4.80	0.00
10	29.71	3465	2-methyltetratriacontane	449 (M-43), 477 (M-15)	5.42	0.03
11	29.98	3492	2, 24-dimethyltetratriacontane	168, 336, 365, 463 (M-43), 491 (M-15)	1.21	11.57
12	30.58	3552	11, 15-dimethylpentatriacontane	168, 211, 239, 280, 308, 379	0.98	12.96
13	30.81	3570	11, 15, 19-trimethylpentatriacontane	168, 239, 252, 309, 323, 393	0.58	13.20
14	31.43	3652	12, 16-; 12, 18-dimethylhexatriacontane	182, 253, 280, 308, 351, 379	0.05	9.59
15	31.68	3669	3-methylhexatriacontane	491 (M-29)	0.10	15.23
16	31.97	3692	4, 10- and 4, 12-dimethylhexatriacontane	168, 197, 225, 336, 364, 393	3.04	6.94
17	32.17	3711	4, 10, 14-trimethylhexatriacontane	168, 239, 336, 407	5.82	0.21
18	32.63	3729	11-, 13-, 15-methylheptatriacontane	168, 196, 224, 280, 308, 336, 364	1.87	8.22
19	33.03	3751	11, 15-, and 11, 19-dimethylheptatriacontane	168, 239, 280, 295, 407	2.44	8.90
20	33.35	3771	11, 15,19-trimethylhepatriacontane	168, 239, 280, 309, 351	6.73	0.17
21	33.61	3789	11, 15, 23-trimethylheptatriacontane	168, 224, 239	5.15	0.84
22	34.15	3837	unknown		6.31	0.14
23	34.48	3849	12, 16-dimethyloctatriacontane	182, 407, 253, 336, 309	6.04	1.21
24	34.84	3868	unknown	182, 225	4.85	0.65
25	36.31	3953	11, 21-dimethylnonatriacontane	168, 280, 295, 323, 407, 435	5.47	0.99
26	36.7	3973	11, 15, 19-trimethylnonatriacontane	168, 239, 309, 379, 449	5.33	0.64
27	37.89	4072	12, 16-dimethyltetracontane	182, 253, 378, 449	3.63	1.12
28	40.85	4186	11, 15, 19-trimethylhentetracontane	168, 239, 280, 308	3.78	0.38
Total					100	100

Table 3. Prediction errors in the classification of three *Macrolophus* species based on a quadratic discriminant analysis (QDA) using the three main principal components of the cuticular hydrocarbons of males and females. The number in each cell represents the number of individuals of the expected species (column) assigned to the three predicted species (rows). A total 119 insects where analysed.

	Expected species			
Predicted species by QDA	*M. melanotoma*	*M. pygmaeus*	*M. costalis*	Total
M. melanotoma	39	0	0	39
M. pygmaeus	0	38	2	40
M. costalis	0	5	35	40
Total	39	43	37	119
Prediction error (%)	0	13.16	5.4	5.88

Fig. 1. Representative gas chromatograph traces of *Macrolophus* cuticular hydrocarbons. Each trace corresponds to a different single individual. The most representative peaks were numbered 1 to 28, corresponding with the column elution time. Notice the large difference between *M. melanotoma* and the other two species as well as between *M. melanotoma* males and females. Arrow indicates the characteristic dark spot on the scutellum of *M. costalis*.

Fig. 2. Relative abundance of male and female cuticular hydrocarbons in three *Macrolophus* species. Abundance is the percentage relative to the internal standard. Asterisks above each bar pair indicate significant difference between males (grey bars) and females (empty bars) (t-test, $P < 0.0017$). Different letters for a given compound indicate significant differences among species in females (capital letters) or males (small letters) (ANOVA followed by Tukey´s test, $P < 0.0017$ in both).

Fig. 3. Effect of diet on *Macrolophus* cuticular hydrocarbons. One group ("field" individuals) fed on their natural diet (white bars), and another group ("laboratory" individuals) were reared in a standard tobacco and moth egg diet (black bars). Abundance is the percentage relative to the internal standard. Asterisks above a bar pair indicate significant difference between diets (t-test, $P < 0.0017$).

Fig. 4. Contribution of the two largest principal components (PC) from a PCA of *Macrolophus* cuticular hydrocarbons. PC1 explains 46 % of the variability and separates *M. melanotoma* males (white circles) from the females (black circles), and from the other two species (*M. pygmaeus*, triangles, *M. costalis*, squares). PC2 explains 22% of the variability and separates *M. melanotoma* females from the males and from

505 the other two species. Laboratory individuals marked with a dot, field individuals

506 unmarked.

Figure 1

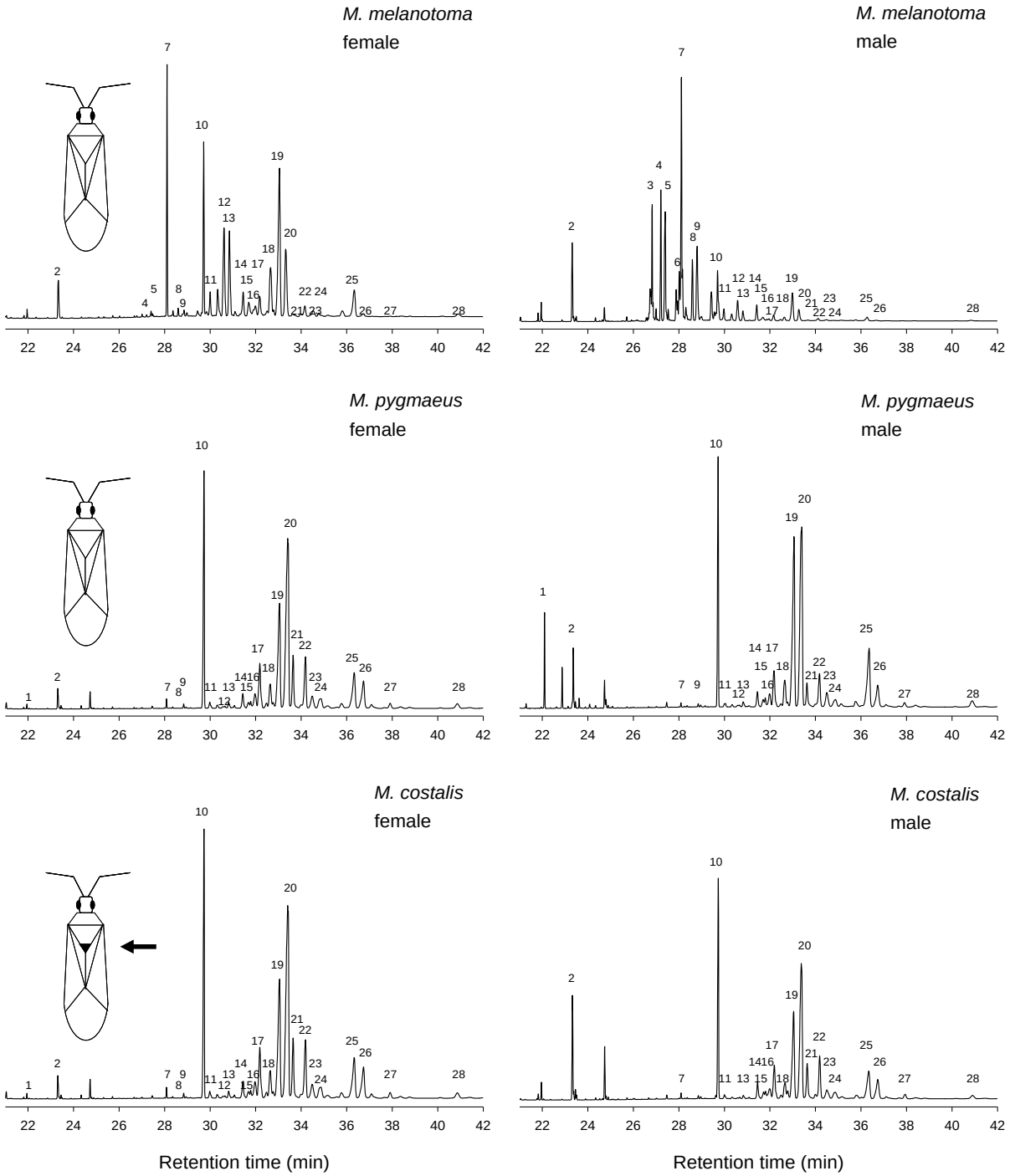

M. melanotoma
female
M. melanotoma
male
M. pygmaeus
female
M. pygmaeus
male
M. costalis
female
M. costalis
male
Retention time (min)
Retention time (min)

Figure 2

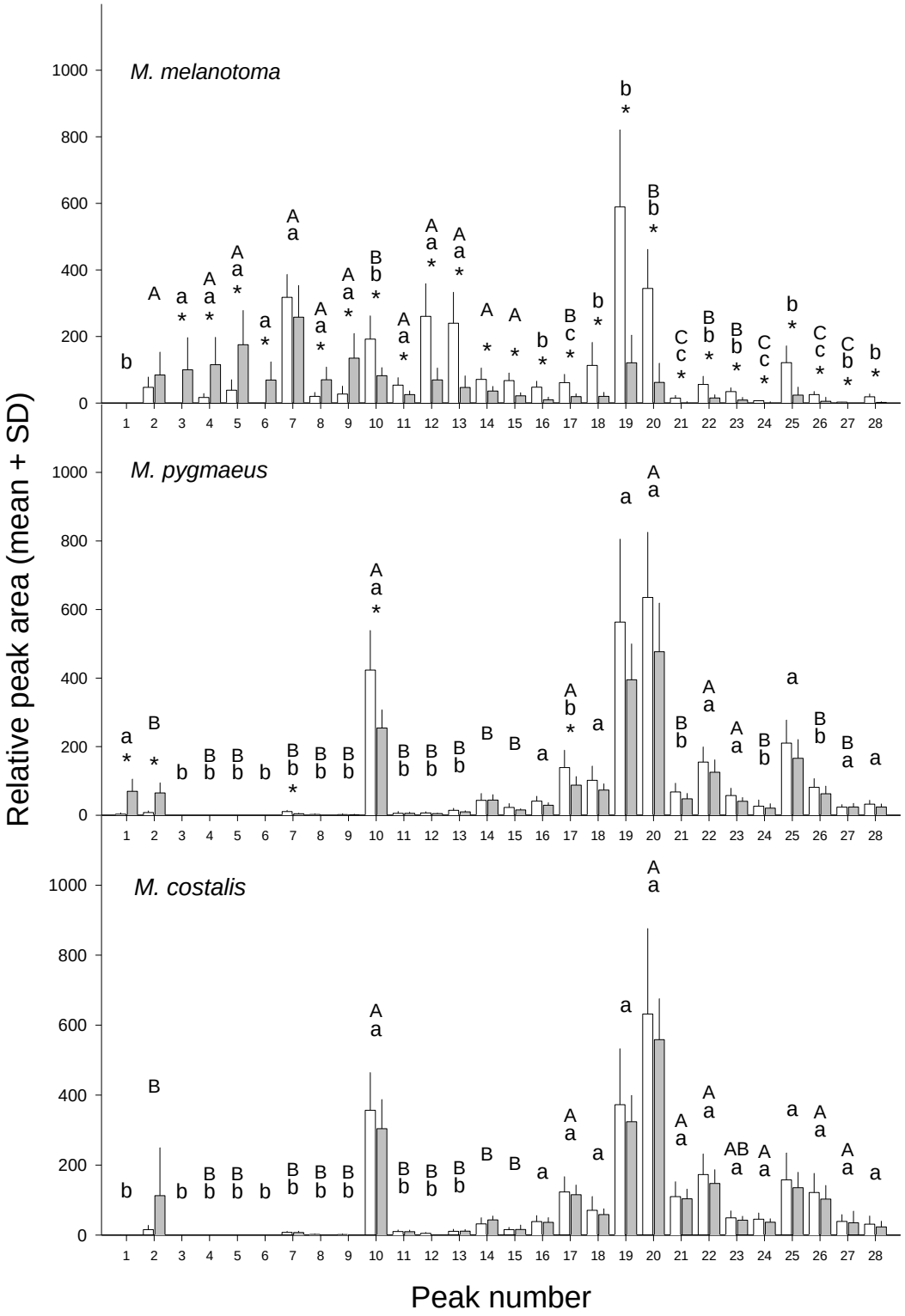

M. melanotoma
M. pygmaeus
M. costalis
Relative peak area (mean + SD)
Peak number

Figure3

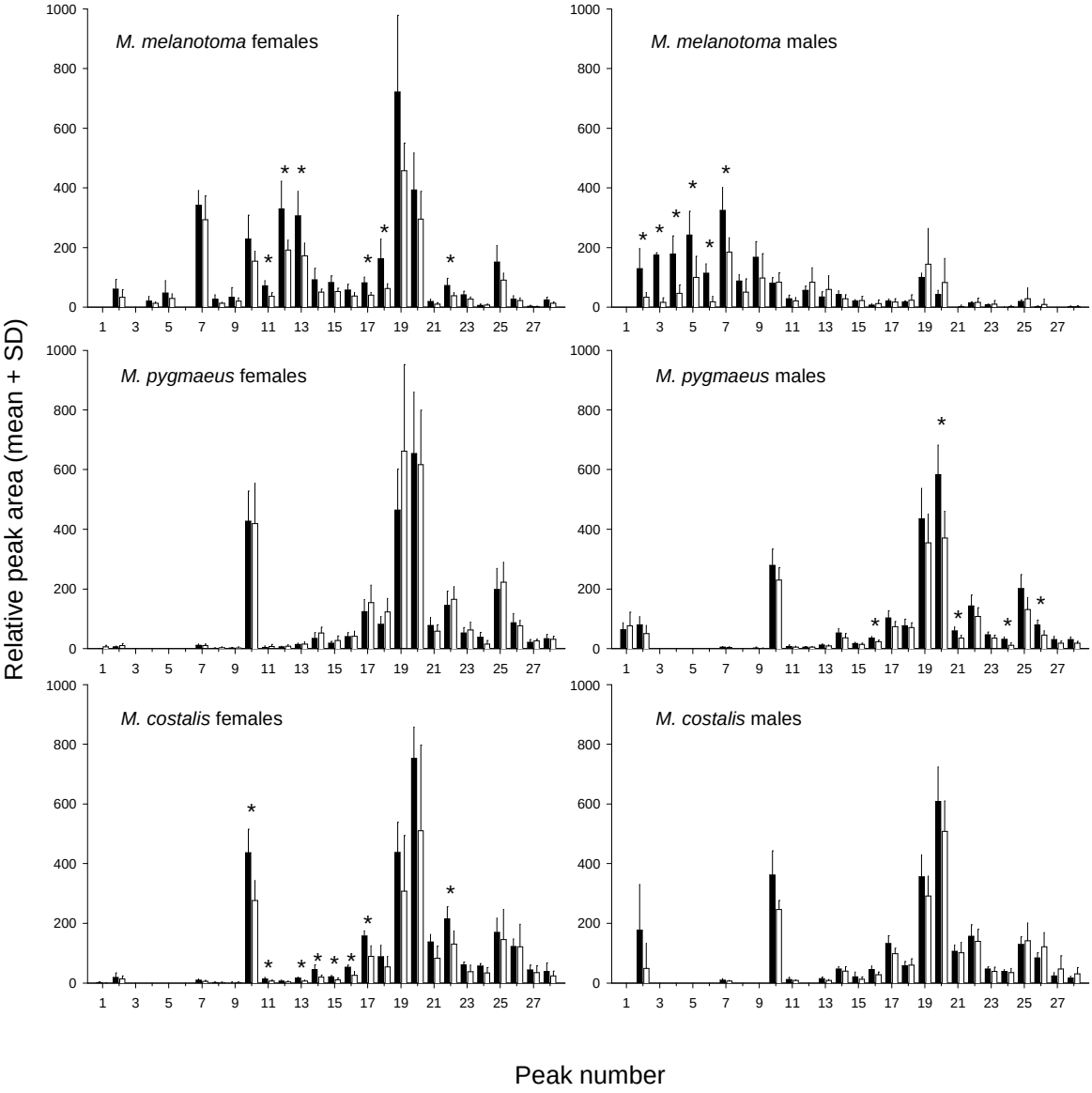
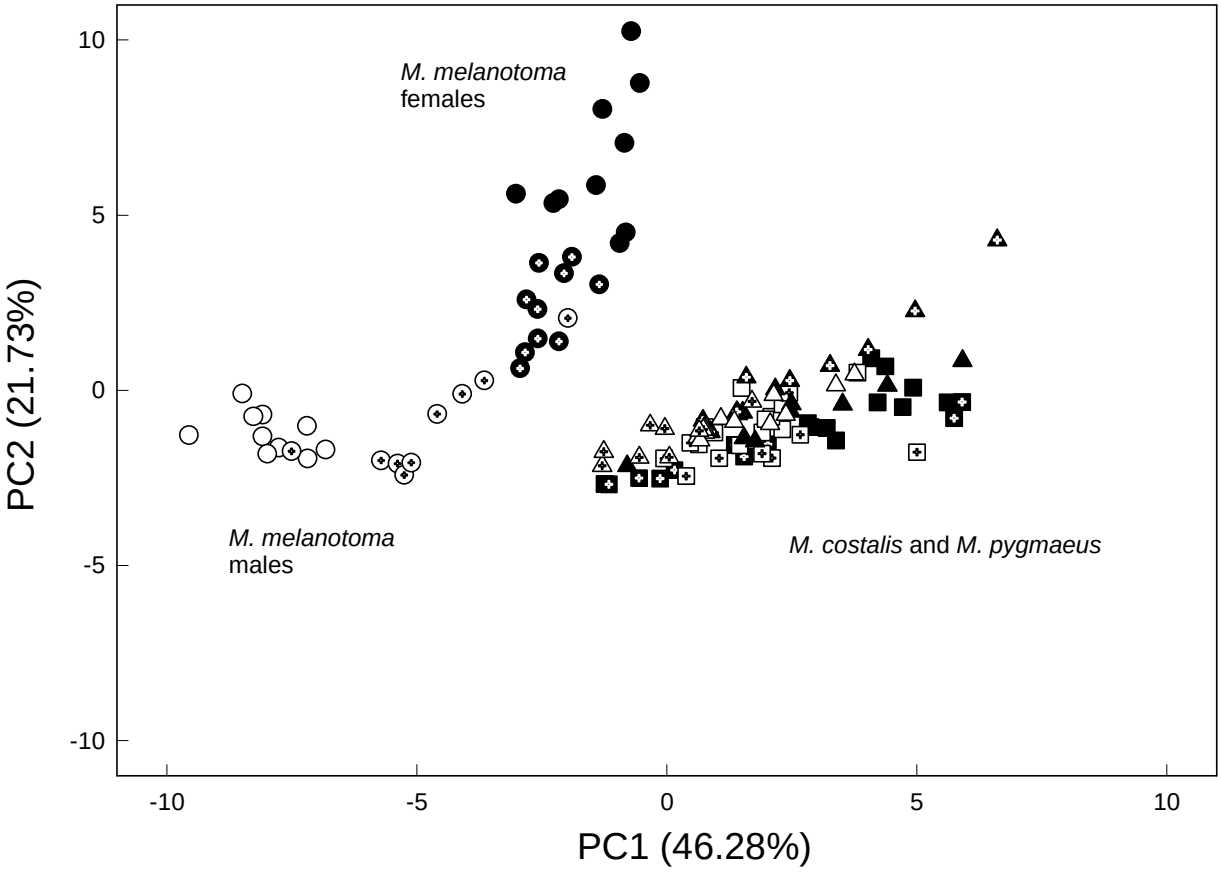

M. melanotoma females
M. melanotoma males
M. pygmaeus females
M. pygmaeus males
M. costalis females
M. costalis males
Relative peak area (mean + SD)
Peak number

Figure4


10
5
0
-5
-10
PC2 (21.73%)
M. melanotoma
females
M. melanotoma
males
M. costalis and M. pygmaeus
-10
-5
0
5
10
PC1 (46.28%)

Supplementary Material for

Cuticular hydrocarbons discriminate isomorphic *Macrolophus* species (Heteroptera: Miridae)

César Gemeno[1*], Nerea Laserna[1], Magí Riba[2], Joan Valls[3], Cristina Castañé[4] and Oscar Alomar[4]

[1] *Department of Crop and Forest Science, University of Lleida, 25198 Lleida, Spain*

[2] *Department of Chemistry, University of Lleida, 25198 Lleida, Spain*

[3] *Biostatistics Unit. Biomedical Research Institute (IRBLLEIDA), 25198 Lleida, Spain*

[4] *IRTA Cabrils, 08348 Barcelona, Spain*

1. Supplementary Tables

Males	True species			
Predicted species by QDA	*M. melanotoma*	*M. pygmaeus*	*M. costalis*	Total
M. melanotoma	19	0	0	19
M. pygmaeus	0	20	0	20
M. costalis	0	1	19	20
Total	19	21	19	59
Prediction error (%)	0	4.76	0	1.7

Table S1. Prediction errors in the classification of *Macrolophus* species based on a quadratic discriminant analysis (QDA) using the three main principal components of the cuticular hydrocarbons of males. The number in each cell represents the number of individuals of a true species (column) assigned to the three predicted species (rows). A total 59 insects was analysed.

Females	True species			
Predicted species by QDA	*M. melanotoma*	*M. pygmaeus*	*M. costalis*	Total
M. melanotoma	20	0	0	20
M. pygmaeus	0	20	1	20
M. costalis	0	0	20	20
Total	20	20	21	60
Prediction error (%)	0	0	4.76	1.7

Table S2. Prediction errors in the classification of *Macrolophus* species based on a quadratic discriminant analysis (QDA) using the three main principal components of the cuticular hydrocarbons of females. The number in each cell represents the number of individuals of a true species (column) assigned to the three predicted species (rows). A total 60 insects was analysed.

2. Functions for *Macrolophus* cuticular hydrocarbon QDA computation with R

Three files are available as supplementary material, as described below:

1) **S2.Macropholus.data.TXT** Contains the relative areas of the 28 cuticular hydrocarbons from 119 insects corresponding to three species: *M. melatonoma* (Me), *M. pygmaeus* (Py), and *M. Costalis* (Co), two sexes (m, f), and two origins: laboratory (L) and field (F), analyzed in the main paper.
2) **S3.Macrolophus.QDA.R** Contains the script (in R code) that allows reproducing the results shown in the main paper.
3) **S4.Macrolophus.predict.R** Contains the code (in R) of the function predict.QDA(), which allows to compute the predicted probabilities for new individuals, using our data as basis.

To reproduce the results shown in the main paper in R, the script (**S3.Macrolophus.QDA.R**) has to be executed using the datafile provided "**S2.Macrolophus.data.TXT**". This script can also be used to predict the species for a set of new individuals. Below we show, step by step, how we obtained the results of the main paper with R code. We also analyze males and females separately. In addition we analyze a male and a female *Macrolophus sp*. collected on tomato plants in Cabrils (Barcelona, Spain), by using the script contained in the file "**S4.Macrolophus.predict.R**".

2.1 First, the data has to be read, using the data file "S2.Macrolophus.data.TXT"

```
> ########################
> #  1 Data reading      #
> ########################

> macrolophus <- read.table("S2.Macropholus.data.TXT",sep="\t",header=T)
> summary(macrolophus)
 species sex   origin       p1               p2               p3               p4
 Me:39   f:60   F:59   Min.   :  0.000   Min.   :  1.63   Min.   :  0.00   Min.   :  0.00
 Co:40   m:59   L:60   1st Qu.:  0.000   1st Qu.: 11.42   1st Qu.:  0.00   1st Qu.:  0.00
 Py:40                 Median :  0.000   Median : 29.20   Median :  0.00   Median :  0.00
                       Mean   : 12.320   Mean   : 55.08   Mean   : 15.91   Mean   : 21.17
                       3rd Qu.:  2.005   3rd Qu.: 74.72   3rd Qu.:  0.00   3rd Qu.: 11.88
                       Max.   :150.710   Max.   :547.65   Max.   :318.30   Max.   :259.51
      p5               p6               p7               p8               p9
 Min.   :  0.00   Min.   :  0.00   Min.   :  0.00   Min.   :  0.00   Min.   :  0.00
 1st Qu.:  0.00   1st Qu.:  0.00   1st Qu.:  5.73   1st Qu.:  0.00   1st Qu.:  0.00
 Median :  0.00   Median :  0.00   Median :  9.29   Median :  1.24   Median :  1.35
 Mean   : 34.44   Mean   : 10.99   Mean   : 99.57   Mean   : 15.14   Mean   : 26.82
 3rd Qu.: 25.82   3rd Qu.:  0.00   3rd Qu.:235.72   3rd Qu.: 14.39   3rd Qu.: 17.43
 Max.   :393.66   Max.   :157.52   Max.   :476.77   Max.   :157.47   Max.   :294.53
     p10              p11              p12              p13              p14
 Min.   : 52.81   Min.   : 0.00   Min.   :  0.00   Min.   :  2.460   Min.   :  0.00
 1st Qu.:172.97   1st Qu.: 4.93   1st Qu.:  2.64   1st Qu.:  9.275   1st Qu.: 30.93
 Median :252.68   Median : 9.60   Median :  5.64   Median : 15.670   Median : 42.01
 Mean   :270.35   Mean   :18.13   Mean   : 57.47   Mean   : 55.422   Mean   : 44.99
 3rd Qu.:355.94   3rd Qu.:23.34   3rd Qu.: 54.49   3rd Qu.: 30.025   3rd Qu.: 55.42
 Max.   :677.64   Max.   :91.32   Max.   :483.56   Max.   :419.810   Max.   :131.52
     p15              p16              p17              p18              p19
 Min.   :  0.00   Min.   : 2.62   Min.   :  4.73   Min.   :  6.86   Min.   :  48.15
 1st Qu.: 13.29   1st Qu.:18.12   1st Qu.: 52.55   1st Qu.: 45.41   1st Qu.: 262.06
 Median : 18.85   Median :32.79   Median : 91.42   Median : 64.00   Median : 377.50
 Mean   : 26.40   Mean   :33.88   Mean   : 91.38   Mean   : 73.33   Mean   : 396.46
 3rd Qu.: 26.71   3rd Qu.:47.68   3rd Qu.:129.88   3rd Qu.: 92.42   3rd Qu.: 479.30
 Max.   :126.30   Max.   :86.88   Max.   :288.27   Max.   :277.26   Max.   :1374.73
     p20              p21              p22              p23              p24
 Min.   :  24.01   Min.   :  0.00   Min.   :  3.92   Min.   :  2.07   Min.   : 0.000
 1st Qu.: 285.81   1st Qu.: 16.29   1st Qu.: 53.40   1st Qu.: 24.36   1st Qu.: 5.505
 Median : 467.93   Median : 49.50   Median :118.32   Median : 40.79   Median :19.320
 Mean   : 454.72   Mean   : 57.84   Mean   :112.74   Mean   : 39.03   Mean   :22.953
 3rd Qu.: 634.28   3rd Qu.: 88.03   3rd Qu.:156.94   3rd Qu.: 50.22   3rd Qu.:36.525
 Max.   :1140.66   Max.   :176.75   Max.   :283.98   Max.   :112.66   Max.   :69.310
     p25              p26              p27              p28
 Min.   :  5.98   Min.   :  0.00   Min.   :  0.00   Min.   :  0.00
```

```
 1st Qu.: 85.47   1st Qu.: 24.89   1st Qu.:  1.70   1st Qu.: 11.93
 Median :131.88   Median : 63.68   Median : 20.09   Median : 20.32
 Mean   :136.66   Mean   : 66.90   Mean   : 20.85   Mean   : 21.96
 3rd Qu.:179.68   3rd Qu.: 96.83   3rd Qu.: 30.48   3rd Qu.: 28.04
 Max.   :354.38   Max.   :273.62   Max.   :167.80   Max.   :117.71
> dim(macrolophus)
[1] 119  31
> head(macrolophus)
  species sex origin p1     p2 p3    p4    p5 p6     p7    p8    p9    p10   p11    p12    p13
1      Me    f      L  0  14.45  0 13.67 35.07  0 283.19 21.36 19.44 147.78 61.97 342.83 272.28
2      Me    f      L  0  26.25  0  4.90 10.70  0 285.29  8.70  7.17 217.82 41.70 227.25 269.16
3      Me    f      L  0  22.27  0  7.21 22.22  0 354.04 12.90 13.83 211.20 83.76 249.88 419.81
4      Me    f      L  0 112.70  0 35.17 66.40  0 291.57 48.99 26.85 136.85 79.55 381.80 341.69
5      Me    f      L  0  83.21  0 41.98 79.83  0 426.06 44.55 61.78 287.90 66.09 422.08 259.76
6      Me    f      L  0  59.55  0 27.08 62.57  0 407.93 25.29 43.05 147.60 91.32 289.20 321.04
     p14    p15   p16    p17    p18     p19    p20   p21   p22   p23   p24    p25   p26  p27   p28
1  99.85  75.50 45.42  57.17 111.53  590.21 311.86  9.65 62.45 32.59  7.02  98.23 16.11 1.70 11.86
2  74.01  66.68 52.13  59.52 136.79  634.74 395.99 28.36 71.38 39.22  6.80 157.56 30.79 4.69 25.41
3  68.35  70.34 76.38  66.94  58.73  428.78 449.71 23.21 33.47 27.32  8.73  76.08 25.21 0.82 14.48
4 128.09 103.03 73.26  88.76 145.39  831.35 474.16 34.49 80.34 49.89 10.00 171.57 29.66 2.92 25.51
5 131.52  97.00 50.55 103.16 277.26 1090.02 429.99  8.62 93.84 54.69  7.25 239.37 48.04 4.74 34.73
6  99.89  64.28 53.26  69.80  88.64  462.01 287.52 18.29 54.26 26.29  5.21  85.66 17.36 1.60 13.18
```

2.2 Next, the code below can be used to perform the principal component analysis.

```
> ########################################
> #  2 Principal components analysis     #
> ########################################

> library(MASS)
> macrolophus.pca <- princomp(macrolophus[,4:31],cor=T)
> summary(macrolophus.pca)
Importance of components:
                          Comp.1    Comp.2    Comp.3     Comp.4     Comp.5     Comp.6     Comp.7
Standard deviation     3.5997770 2.4667929 1.6782470 1.14622797 1.04554207 0.89193104 0.72343036
Proportion of Variance 0.4627998 0.2173238 0.1005897 0.04692281 0.03904136 0.02841218 0.01869112
Cumulative Proportion  0.4627998 0.6801236 0.7807134 0.82763620 0.86667756 0.89508974 0.91378086
                           Comp.8     Comp.9   Comp.10    Comp.11     Comp.12     Comp.13
Standard deviation     0.66406272 0.65833720 0.6079244 0.46722859 0.426682075 0.365307627
Proportion of Variance 0.01574926 0.01547885 0.0131990 0.00779652 0.006502057 0.004766059
Cumulative Proportion  0.92953012 0.94500898 0.9582080 0.96600450 0.972506554 0.977272613
                           Comp.14     Comp.15     Comp.16     Comp.17     Comp.18     Comp.19
Standard deviation     0.336951113 0.301806189 0.287428171 0.267444218 0.251891725 0.228047493
Proportion of Variance 0.004054859 0.003253106 0.002950534 0.002554515 0.002266051 0.001857345
Cumulative Proportion  0.981327472 0.984580579 0.987531113 0.990085627 0.992351679 0.994209024
                           Comp.20     Comp.21      Comp.22      Comp.23      Comp.24      Comp.25
Standard deviation     0.200466700 0.188514348 0.1508854859 0.1484273877 0.1166391642 0.0933369246
Proportion of Variance 0.001435246 0.001269202 0.0008130868 0.0007868103 0.0004858820 0.0003111351
Cumulative Proportion  0.995644270 0.996913472 0.9977265590 0.9985133694 0.9989992513 0.9993103864
                            Comp.26      Comp.27      Comp.28
Standard deviation     0.0920458179 0.0809715804 0.0654244002
Proportion of Variance 0.0003025869 0.0002341570 0.0001528697
Cumulative Proportion  0.9996129733 0.9998471303 1.0000000000
> macrolophus.pca
Call:
princomp(x = macrolophus[, 4:31], cor = T)

Standard deviations:
    Comp.1     Comp.2     Comp.3     Comp.4     Comp.5     Comp.6     Comp.7     Comp.8     Comp.9
3.59977704 2.46679290 1.67824699 1.14622797 1.04554207 0.89193104 0.72343036 0.66406272 0.65833720
   Comp.10    Comp.11    Comp.12    Comp.13    Comp.14    Comp.15    Comp.16    Comp.17    Comp.18
0.60792436 0.46722859 0.42668208 0.36530763 0.33695111 0.30180619 0.28742817 0.26744422 0.25189173
   Comp.19    Comp.20    Comp.21    Comp.22    Comp.23    Comp.24    Comp.25    Comp.26    Comp.27
0.22804749 0.20046670 0.18851435 0.15088549 0.14842739 0.11663916 0.09333692 0.09204582 0.08097158
   Comp.28
0.06542440

 28  variables and  119 observations.
```

2.3 Quadratic discriminant analysis is then applied to the first three principal components.

```
> ########################################
> #  3 Quadratic discriminant analysis   #
> ########################################

> macrolophus$PC1 <- macrolophus.pca$scores[,1]
> macrolophus$PC2 <- macrolophus.pca$scores[,2]
> macrolophus$PC3 <- macrolophus.pca$scores[,3]
> macrolophus.qda <- qda(species~PC1+PC2+PC3,data=macrolophus)
> macrolophus.qda
Call:
qda(species ~ PC1 + PC2 + PC3, data = macrolophus)
```

```
Prior probabilities of groups:
       Me        Co        Py
0.3277311 0.3361345 0.3361345

Group means:
         PC1         PC2         PC3
Me -4.115516 -1.7241258 -0.01944857
Co  2.238575  1.1644365 -0.31066728
Py  1.774053  0.5165862  0.32962964
>
> macrolophus.predict <- predict(macrolophus.qda)
> macrolophus.predict
$class
  [1] Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me Me
 [33] Me Me Me Me Me Me Me Py Co Py Co Co Co Co Co Co Co Co Co Co Py Co Co Co Co Co Co Co Co Co
 [65] Co Co Co Co Co Co Co Co Co Py Co Co Co Py Co Py Co Co Py Py Py Py Py Py Py Py Py Py Py
 [97] Py Py Py Py Py Py Py Py Py Py Py Py Py Py Py Py Py Py Py Py Py Py Py Py
Levels: Me Co Py

$posterior
              Me             Co             Py
1   1.000000e+00  1.081900e-35  2.972536e-64
2   1.000000e+00  1.232411e-26  2.455322e-31
3   1.000000e+00  1.219622e-37  6.922950e-56
4   1.000000e+00  1.459947e-56 1.182524e-135
5   1.000000e+00  3.224060e-61 2.643045e-155
6   1.000000e+00  1.404550e-38 6.018271e-112
7   1.000000e+00  9.535992e-47  1.504535e-76
8   1.000000e+00  6.277640e-80 1.994038e-201
9   1.000000e+00  2.604715e-22  2.265620e-37
10  1.000000e+00  1.266279e-38  6.375300e-58
11  1.000000e+00  1.221547e-22  2.325122e-46
12  1.000000e+00  1.359212e-09  1.243514e-19
13  1.000000e+00  6.943087e-23  1.504406e-64
14  1.000000e+00  2.021508e-09  4.584161e-30
15  1.000000e+00  1.496023e-15  1.984849e-30
16  1.000000e+00  1.625087e-11  8.347107e-21
17  1.000000e+00  2.260344e-15  5.050571e-33
18  1.000000e+00  5.379995e-17  1.082916e-47
19  1.000000e+00  8.509437e-20  1.411362e-44
20  9.999999e-01  5.821288e-08  2.213222e-20
21  1.000000e+00 2.072800e-230  0.000000e+00
22  1.000000e+00 1.661224e-295  0.000000e+00
23  1.000000e+00 7.884352e-247  0.000000e+00
24  1.000000e+00 5.709844e-226  0.000000e+00
25  1.000000e+00 1.638139e-237  0.000000e+00
26  1.000000e+00 8.049419e-125  0.000000e+00
27  1.000000e+00 8.863101e-139  0.000000e+00
28  1.000000e+00  0.000000e+00  0.000000e+00
29  1.000000e+00 5.620462e-163  0.000000e+00
30  1.000000e+00 2.198491e-244  0.000000e+00
31  1.000000e+00  6.789367e-47 3.284587e-294
32  9.999998e-01  1.514082e-07  4.015008e-27
33  9.999999e-01  1.215987e-07  5.691264e-45
34  1.000000e+00  3.695892e-15 4.534817e-112
35  1.000000e+00  6.023561e-33 2.339525e-215
36  1.000000e+00  2.402909e-21 1.330354e-146
37  1.000000e+00 6.972786e-149  0.000000e+00
38  1.000000e+00  1.685027e-11  3.405352e-45
39  1.000000e+00  5.076884e-19 3.514107e-135
40  4.629291e-60  3.084305e-01  6.915695e-01
41  7.906338e-74  9.707091e-01  2.929088e-02
42  2.302682e-54  2.825718e-01  7.174282e-01
43  9.826798e-64  9.549499e-01  4.505011e-02
44  6.819931e-88  9.978319e-01  2.168099e-03
45  4.315915e-73  9.913138e-01  8.686161e-03
46  9.856785e-47  9.459615e-01  5.403853e-02
47  1.129015e-55  9.900679e-01  9.932139e-03
48  1.502124e-46  9.958254e-01  4.174638e-03
49  8.083689e-52  9.786686e-01  2.133140e-02
50  5.824502e-94  9.999187e-01  8.129419e-05
51  1.331753e-29  8.271217e-01  1.728783e-01
52  8.399060e-18  7.802386e-01  2.197614e-01
53  3.949136e-94  9.993729e-01  6.271310e-04
54  6.198120e-10  4.869743e-01  5.130257e-01
55  4.098025e-33  9.511308e-01  4.886918e-02
56  5.317475e-37  8.368314e-01  1.631686e-01
57  1.345028e-19  7.032355e-01  2.967645e-01
58  2.585174e-14  5.899638e-01  4.100362e-01
59  1.449888e-10  7.352459e-01  2.647541e-01
60  2.613829e-41  9.999492e-01  5.075336e-05
61  5.716691e-37  9.999980e-01  1.962613e-06
62  8.326286e-32  9.974282e-01  2.571838e-03
63  2.917459e-35  8.904013e-01  1.095987e-01
64  9.361567e-37  9.584805e-01  4.151953e-02
65  2.381528e-53  9.432295e-01  5.677047e-02
66  2.693560e-22  9.197214e-01  8.027859e-02
```

```
67  2.029885e-27  1.000000e+00  5.895113e-15
68  6.929522e-35  9.952248e-01  4.775237e-03
69  3.708216e-24  1.000000e+00  2.527683e-30
70  1.437963e-88  1.000000e+00  1.754505e-08
71  2.382960e-36  7.018986e-01  2.981014e-01
72  1.256957e-40  9.946113e-01  5.388716e-03
73  1.705737e-22  1.000000e+00  3.943261e-11
74  8.435407e-17  4.496036e-01  5.503964e-01
75  9.077042e-44  7.998866e-01  2.001134e-01
76  4.778201e-38  9.981186e-01  1.881379e-03
77  7.973526e-28  9.566672e-01  4.333277e-02
78  1.025495e-23  4.383259e-01  5.616741e-01
79  2.659344e-22  8.470499e-01  1.529501e-01
80  3.705368e-38  1.769242e-01  8.230758e-01
81  1.010249e-52  8.249005e-01  1.750995e-01
82  5.759876e-85  5.612043e-01  4.387957e-01
83  9.417009e-27  2.606889e-02  9.739311e-01
84  2.647101e-37  7.191369e-02  9.280863e-01
85  8.396146e-62  3.318818e-01  6.681182e-01
86  6.991704e-32  3.233243e-02  9.676676e-01
87  1.114155e-11  1.428729e-01  8.571271e-01
88  4.022769e-32  3.385770e-01  6.614230e-01
89  4.779310e-30  3.028359e-01  6.971641e-01
90  3.111415e-42  3.854323e-03  9.961457e-01
91  1.171852e-23  1.657259e-02  9.834274e-01
92  5.539828e-50  1.182253e-03  9.988177e-01
93  9.494014e-20  2.341961e-02  9.765804e-01
94  9.529373e-59  4.491360e-05  9.999551e-01
95  4.270801e-72  4.650132e-09  1.000000e+00
96  1.803989e-26  1.047406e-01  8.952594e-01
97  4.936870e-34  8.475899e-03  9.915241e-01
98  1.408292e-25  2.692238e-02  9.730776e-01
99  3.773500e-22  4.667346e-02  9.533265e-01
100 1.117377e-19  3.564483e-02  9.643552e-01
101 2.813980e-32  5.480513e-02  9.451949e-01
102 3.803048e-37  1.518908e-01  8.481092e-01
103 1.613566e-34  2.271748e-01  7.728252e-01
104 4.896069e-50  1.984771e-02  9.801523e-01
105 5.417494e-47  6.799428e-02  9.320057e-01
106 1.376361e-21  1.002046e-01  8.997954e-01
107 1.551351e-23  3.468337e-01  6.531663e-01
108 3.035439e-21  9.710385e-02  9.028962e-01
109 1.807461e-26  1.511359e-01  8.488641e-01
110 1.650914e-11  3.619455e-02  9.638055e-01
111 2.187144e-08  4.318655e-02  9.568134e-01
112 1.198214e-07  4.263977e-02  9.573601e-01
113 6.485381e-20  2.052001e-02  9.794800e-01
114 3.303492e-14  1.771733e-01  8.228267e-01
115 2.671738e-11  3.316938e-02  9.668306e-01
116 6.112121e-18  4.239882e-02  9.576012e-01
117 4.653508e-14  2.692003e-02  9.730800e-01
118 1.015974e-19  2.691500e-02  9.730850e-01
119 1.355224e-25  8.503637e-03  9.914964e-01
```

```
> macrolophus$QDA.prob.Me <- macrolophus.predict$posterior[,1]
> macrolophus$QDA.prob.Co <- macrolophus.predict$posterior[,2]
> macrolophus$QDA.prob.Py <- macrolophus.predict$posterior[,3]
> macrolophus$QDA.predicted.species <- macrolophus.predict$class
```

The following data frame contains the predictions for the insects in the dataset.

```
> results <- macrolophus[,c("species","QDA.prob.Me","QDA.prob.Co","QDA.prob.Py","QDA.predicted.species")]
> results[,2:4] <- round(results[,2:4],2)
> results
    species QDA.prob.Me QDA.prob.Co QDA.prob.Py QDA.predicted.species
1        Me           1        0.00        0.00                    Me
2        Me           1        0.00        0.00                    Me
3        Me           1        0.00        0.00                    Me
4        Me           1        0.00        0.00                    Me
5        Me           1        0.00        0.00                    Me
6        Me           1        0.00        0.00                    Me
7        Me           1        0.00        0.00                    Me
8        Me           1        0.00        0.00                    Me
9        Me           1        0.00        0.00                    Me
10       Me           1        0.00        0.00                    Me
11       Me           1        0.00        0.00                    Me
12       Me           1        0.00        0.00                    Me
13       Me           1        0.00        0.00                    Me
14       Me           1        0.00        0.00                    Me
15       Me           1        0.00        0.00                    Me
16       Me           1        0.00        0.00                    Me
17       Me           1        0.00        0.00                    Me
18       Me           1        0.00        0.00                    Me
19       Me           1        0.00        0.00                    Me
20       Me           1        0.00        0.00                    Me
21       Me           1        0.00        0.00                    Me
22       Me           1        0.00        0.00                    Me
```

23	Me	1	0.00	0.00	Me
24	Me	1	0.00	0.00	Me
25	Me	1	0.00	0.00	Me
26	Me	1	0.00	0.00	Me
27	Me	1	0.00	0.00	Me
28	Me	1	0.00	0.00	Me
29	Me	1	0.00	0.00	Me
30	Me	1	0.00	0.00	Me
31	Me	1	0.00	0.00	Me
32	Me	1	0.00	0.00	Me
33	Me	1	0.00	0.00	Me
34	Me	1	0.00	0.00	Me
35	Me	1	0.00	0.00	Me
36	Me	1	0.00	0.00	Me
37	Me	1	0.00	0.00	Me
38	Me	1	0.00	0.00	Me
39	Me	1	0.00	0.00	Me
40	Co	0	0.31	0.69	Py
41	Co	0	0.97	0.03	Co
42	Co	0	0.28	0.72	Py
43	Co	0	0.95	0.05	Co
44	Co	0	1.00	0.00	Co
45	Co	0	0.99	0.01	Co
46	Co	0	0.95	0.05	Co
47	Co	0	0.99	0.01	Co
48	Co	0	1.00	0.00	Co
49	Co	0	0.98	0.02	Co
50	Co	0	1.00	0.00	Co
51	Co	0	0.83	0.17	Co
52	Co	0	0.78	0.22	Co
53	Co	0	1.00	0.00	Co
54	Co	0	0.49	0.51	Py
55	Co	0	0.95	0.05	Co
56	Co	0	0.84	0.16	Co
57	Co	0	0.70	0.30	Co
58	Co	0	0.59	0.41	Co
59	Co	0	0.74	0.26	Co
60	Co	0	1.00	0.00	Co
61	Co	0	1.00	0.00	Co
62	Co	0	1.00	0.00	Co
63	Co	0	0.89	0.11	Co
64	Co	0	0.96	0.04	Co
65	Co	0	0.94	0.06	Co
66	Co	0	0.92	0.08	Co
67	Co	0	1.00	0.00	Co
68	Co	0	1.00	0.00	Co
69	Co	0	1.00	0.00	Co
70	Co	0	1.00	0.00	Co
71	Co	0	0.70	0.30	Co
72	Co	0	0.99	0.01	Co
73	Co	0	1.00	0.00	Co
74	Co	0	0.45	0.55	Py
75	Co	0	0.80	0.20	Co
76	Co	0	1.00	0.00	Co
77	Co	0	0.96	0.04	Co
78	Co	0	0.44	0.56	Py
79	Co	0	0.85	0.15	Co
80	Py	0	0.18	0.82	Py
81	Py	0	0.82	0.18	Co
82	Py	0	0.56	0.44	Co
83	Py	0	0.03	0.97	Py
84	Py	0	0.07	0.93	Py
85	Py	0	0.33	0.67	Py
86	Py	0	0.03	0.97	Py
87	Py	0	0.14	0.86	Py
88	Py	0	0.34	0.66	Py
89	Py	0	0.30	0.70	Py
90	Py	0	0.00	1.00	Py
91	Py	0	0.02	0.98	Py
92	Py	0	0.00	1.00	Py
93	Py	0	0.02	0.98	Py
94	Py	0	0.00	1.00	Py
95	Py	0	0.00	1.00	Py
96	Py	0	0.10	0.90	Py
97	Py	0	0.01	0.99	Py
98	Py	0	0.03	0.97	Py
99	Py	0	0.05	0.95	Py
100	Py	0	0.04	0.96	Py
101	Py	0	0.05	0.95	Py
102	Py	0	0.15	0.85	Py
103	Py	0	0.23	0.77	Py
104	Py	0	0.02	0.98	Py
105	Py	0	0.07	0.93	Py
106	Py	0	0.10	0.90	Py
107	Py	0	0.35	0.65	Py
108	Py	0	0.10	0.90	Py
109	Py	0	0.15	0.85	Py
110	Py	0	0.04	0.96	Py

```
111      Py           0        0.04        0.96                   Py
112      Py           0        0.04        0.96                   Py
113      Py           0        0.02        0.98                   Py
114      Py           0        0.18        0.82                   Py
115      Py           0        0.03        0.97                   Py
116      Py           0        0.04        0.96                   Py
117      Py           0        0.03        0.97                   Py
118      Py           0        0.03        0.97                   Py
119      Py           0        0.01        0.99                   Py
```

The table below compares the true species with the predicted ones. An error of 5.88% is obtained.

```
>                                          correct.prediction                                          <-
table(macrolophus$species,macrolophus$QDA.predicted.species,dnn=c("real","predicted"))
> correct.prediction
    predicted
real Me Co Py
  Me 39  0  0
  Co  0 35  5
  Py  0  2 38
> prop.table(correct.prediction,1)
    predicted
real    Me    Co    Py
  Me 1.000 0.000 0.000
  Co 0.000 0.875 0.125
  Py 0.000 0.050 0.950
> prop.table(correct.prediction,2)
    predicted
real         Me         Co         Py
  Me 1.00000000 0.00000000 0.00000000
  Co 0.00000000 0.94594595 0.11627907
  Py 0.00000000 0.05405405 0.88372093
>
> total.correct <- sum(diag(correct.prediction))/sum(correct.prediction)
> total.correct
[1] 0.9411765

> total.error <- 1-total.correct
> total.error
[1] 0.05882353
```

2.4 Finally, sections 5.1 and 5.2 in the script ("S3.Macrolophus.QDA.R") contain the code for performing the same analysis above but for each sex individually. The result is shown in supplementary tables 1 and 2 above.

2.5 Here we illustrate the prediction for two *Macrolophus sp.* individuals, one of each sex, of unknown species status, which were collected on tomato plants ("tomatomale" and "tomatofemale").

```
> ########################################
> #  4 Prediction for new individuals    #
> ########################################

> library(mnormt)
> tomatomale <- c(9.3359225,35.24040853,0,0,0,0,2.620230036,4.143500323,0,227.0959237,6.803160605,
1.675141163,5.174382195,27.78788589,33.50457571,21.56733939,80.30059764,62.56136815,290.5071812,451.0842287,
81.21811111,161.8425264,35.04168524,11.66117306,149.6672654,57.47777861,23.24498038,20.68211748)
> tomatofemale <- c(0.465840529,33.08057566,0,1.69478331,0,2.434254523,13.35527135,3.219487532,0.805361245,
623.4437673,21.48072989,5.886884465,11.75026023,17.20756204,32.65718085,44.08958108,133.3946882,100.2624269,6
68.2018557,775.8341467,86.4138592,143.781386,65.40976156,43.59786875,237.3767564,101.9367407,24.22607705,38.8
2679428)
> macrolophus.new <- rbind(tomatomale,tomatofemale)
> source("S4.Macrolophus.predict.R")
> predict.QDA(macrolophus.pca,macrolophus.new,macrolophus,macrolophus$species)
             prob.Ca.cond.X.star prob.Co.cond.X.star prob.Py.cond.X.star
tomatomale                     0               0.335               0.665
tomatofemale                   0               0.074               0.926
```

The results clearly show that the female can be assigned to *M. pygmaeus* (with a probability of 0.92). However, the prediction for the male is less clear. Although it is not a *M. melatonoma* (probability 0), it could be either a *M. costalis* (probability 0.075) or *M. pygmaeus* (probability 0.665). Because the male lacked the characteristic dark spot on the scutellum, it was classified as *M. pygmaeus*. So the model demonstrated that the two insects collected on tomato belong to this last species.

species	sex	origin	p1	p2	p3	p4	p5	p6	p7	p8	p9	p10
	p11	p12	p13	p14	p15	p16	p17	p18	p19	p20	p21	p22
	p23	p24	p25	p26	p27	p28						
Me	f	L	0.00	14.45	0.00	13.67	35.07	0.00	283.19	21.36	19.44	
	147.78	61.97	342.83	272.28	99.85	75.50	45.42	57.17	111.53	590.21	311.86	9.65
	62.45	32.59	7.02	98.23	16.11	1.70	11.86					
Me	f	L	0.00	26.25	0.00	4.90	10.70	0.00	285.29	8.70	7.17	
	217.82	41.70	227.25	269.16	74.01	66.68	52.13	59.52	136.79	634.74	395.99	28.36
	71.38	39.22	6.80	157.56	30.79	4.69	25.41					
Me	f	L	0.00	22.27	0.00	7.21	22.22	0.00	354.04	12.90	13.83	
	211.20	83.76	249.88	419.81	68.35	70.34	76.38	66.94	58.73	428.78	449.71	23.21
	33.47	27.32	8.73	76.08	25.21	0.82	14.48					
Me	f	L	0.00	112.70	0.00	35.17	66.40	0.00	291.57	48.99	26.85	
	136.85	79.55	381.80	341.69	128.09	103.03	73.26	88.76	145.39	831.35	474.16	34.49
	80.34	49.89	10.00	171.57	29.66	2.92	25.51					
Me	f	L	0.00	83.21	0.00	41.98	79.83	0.00	426.06	44.55	61.78	
	287.90	66.09	422.08	259.76	131.52	97.00	50.55	103.16	277.26	1090.02		
	429.99	8.62	93.84	54.69	7.25	239.37	48.04	4.74	34.73			
Me	f	L	0.00	59.55	0.00	27.08	62.57	0.00	407.93	25.29	43.05	
	147.60	91.32	289.20	321.04	99.89	64.28	53.26	69.80	88.64	462.01	287.52	18.29
	54.26	26.29	5.21	85.66	17.36	1.60	13.18					
Me	f	L	0.00	78.24	0.00	11.94	29.35	0.00	350.30	34.72	63.41	
	328.09	87.65	403.34	358.25	0.00	99.90	71.20	106.12	195.55	887.85	475.61	23.58
	89.63	49.39	7.49	175.05	33.25	4.30	22.37					
Me	f	L	0.00	60.74	0.00	48.21	143.60	0.00	304.93	40.71	97.21	
	171.24	88.38	483.56	412.75	121.16	126.30	86.88	107.93	192.98	1154.90		
	624.05	17.25	96.79	59.61	11.09	225.98	41.94	4.61	39.05			
Me	f	L	0.00	59.74	0.00	5.96	12.66	0.00	356.84	10.31	0.00	
	311.84	45.79	192.94	150.29	107.10	51.43	26.15	68.99	207.29	514.11	197.88	9.13
	102.20	36.93	0.00	142.65	20.27	8.66	35.32					
Me	f	L	0.00	89.47	0.00	4.47	12.12	0.00	353.91	19.62	6.70	
	333.17	66.03	303.83	264.59	84.85	70.81	40.99	84.37	215.15	619.46	283.57	19.14
	44.50	38.28	0.00	139.39	15.95	0.00	18.66					
Me	f	F	0.00	45.78	0.00	17.73	27.74	0.00	361.18	16.36	19.56	
	195.88	46.49	214.43	209.15	55.64	68.24	51.32	46.49	80.08	621.04	419.46	13.72
	37.60	32.72	6.40	122.97	29.12	1.52	16.67					
Me	f	F	0.00	13.92	0.00	9.04	17.65	0.00	189.65	11.71	15.27	
	128.94	31.78	150.95	151.25	42.47	44.00	33.94	40.18	59.86	373.44	284.39	16.97
	38.91	29.57	10.27	85.28	24.57	1.78	10.35					
Me	f	F	0.00	21.06	0.00	19.55	63.33	0.00	271.98	16.49	46.79	
	136.53	62.16	245.48	229.28	64.45	65.08	48.01	55.16	52.58	460.07	329.52	12.01
	50.00	34.73	11.92	81.96	25.73	1.70	9.48					
Me	f	F	0.00	28.40	0.00	17.82	35.95	0.00	246.09	14.72	17.74	
	121.76	28.35	144.47	132.55	41.50	40.02	25.73	26.12	52.34	368.55	229.88	7.47
	28.22	19.44	4.06	81.04	17.30	1.44	14.59					
Me	f	F	0.00	13.33	0.00	14.28	30.97	0.00	249.04	14.80	23.11	
	157.94	35.64	178.32	212.80	55.70	61.94	54.72	49.77	64.82	534.58	461.10	19.99
	54.09	37.29	11.29	123.75	41.00	5.89	20.76					
Me	f	F	0.00	14.26	0.00	10.33	19.35	0.00	235.02	8.86	12.66	
	119.60	23.25	184.30	162.62	33.03	52.67	36.17	32.81	44.52	398.05	254.53	6.27
	27.25	22.48	6.04	72.24	19.44	1.54	11.19					

Me	f	F	0.00	13.07	0.00	12.93	32.10	0.00	258.68	13.22	18.93	
	125.76	36.98	211.37	173.17	48.78	55.95	38.49	40.00	48.93	461.17	289.80	5.95
	31.71	24.20	5.80	78.10	22.88	1.46	12.00					
Me	f	F	0.00	30.29	0.00	15.38	37.20	0.00	380.60	16.64	27.48	
	172.18	35.46	194.73	190.71	59.51	48.72	35.41	36.82	71.50	440.15	266.52	8.47
	37.40	21.82	3.82	85.20	15.67	1.50	10.74					
Me	f	F	0.00	75.10	0.00	8.57	23.90	0.00	451.20	12.42	12.53	
	212.83	38.53	229.36	175.35	61.12	52.51	35.77	47.29	94.49	575.30	287.22	5.61
	52.71	34.67	9.62	119.64	23.45	1.95	17.43					
Me	f	F	0.00	77.92	0.00	3.68	7.36	0.00	295.29	5.17	5.66	
	173.75	22.16	158.78	86.07	42.58	32.24	15.05	31.83	59.40	341.88	134.97	0.00
	28.90	17.07	3.02	61.22	9.14	1.12	8.76					
Me	m	L	0.00	131.25	189.04	177.01	281.07	80.71	306.90	90.25	193.02	89.89
	51.46	70.39	49.25	36.37	23.84	10.46	31.25	14.23	106.41	50.89	0.00	9.11
	5.84	0.00	17.22	0.00	0.00	4.27						
Me	m	L	0.00	96.12	178.61	151.41	334.51	157.52	367.91	75.48	236.28	96.12
	42.79	82.85	75.53	68.57	31.58	14.93	28.15	15.14	113.42	71.39	0.00	20.53
	11.00	0.00	20.23	6.41	0.00	0.00						
Me	m	L	0.00	171.27	172.28	200.78	235.16	97.48	289.94	109.00	180.10	69.31
	24.65	52.82	29.46	46.90	20.51	6.26	17.55	16.60	79.93	37.90	0.00	14.48
	7.83	0.00	24.04	0.00	0.00	5.03						
Me	m	L	0.00	117.25	211.30	259.51	156.99	105.95	288.27	96.89	90.89	97.27
	16.76	37.49	20.19	48.04	18.85	4.18	16.82	26.19	100.70	28.07	0.00	17.19
	6.05	0.00	19.44	0.00	0.00	0.00						
Me	m	L	0.00	126.69	143.50	190.06	239.21	82.46	476.77	119.53	200.88	92.99
	36.04	62.10	28.93	40.67	21.14	7.36	21.72	16.12	96.88	40.92	0.00	12.76
	8.38	1.75	17.88	0.00	0.00	3.12						
Me	m	L	0.00	43.69	160.06	134.80	216.29	89.29	274.29	46.36	98.94	57.47
	17.22	43.54	21.26	35.05	20.00	5.30	14.95	15.71	108.28	33.18	0.00	13.48
	6.67	0.00	16.36	0.00	0.00	2.53						
Me	m	L	0.00	77.15	114.72	123.47	158.48	101.55	250.81	90.32	151.90	74.44
	17.66	47.79	19.36	25.79	17.04	3.18	12.39	10.61	81.02	29.67	0.00	6.82
	4.49	0.00	11.85	0.00	0.00	0.00						
Me	m	L	0.00	116.81	318.30	245.20	393.66	155.95	421.11	84.12	203.80	55.68
	25.25	56.17	37.21	40.08	24.59	8.88	22.99	21.50	122.93	60.64	0.00	17.14
	9.21	1.10	29.38	6.73	0.00	0.00						
Me	m	L	0.00	295.56	35.67	58.80	154.98	154.15	240.99	98.38	214.04	
	109.43	34.47	67.82	36.42	52.74	17.36	6.89	25.29	15.49	95.56	39.83	0.00
	16.65	9.76	0.00	19.60	0.00	0.00	0.00					
Me	m	L	0.00	116.79	224.29	231.85	256.51	117.33	329.81	67.69	105.49	63.70
	16.87	42.53	27.97	34.38	17.02	5.69	16.94	16.40	91.32	34.00	0.00	14.18
	7.53	1.42	17.90	0.00	0.00	3.69						
Me	m	F	0.00	69.94	45.77	66.94	119.85	45.34	180.34	60.61	116.66	65.12
	10.08	47.78	26.24	19.60	16.78	5.39	9.08	14.40	76.96	40.01	0.00	7.58
	4.63	0.00	17.09	5.07	0.00	0.00						
Me	m	F	0.00	41.84	4.95	12.95	20.65	0.00	181.64	14.07	17.11	
	102.41	33.64	148.21	116.11	37.22	31.93	17.74	22.98	36.01	228.81	116.19	2.46
	20.27	9.87	1.17	29.64	4.62	0.00	2.71					
Me	m	F	0.00	30.64	3.27	11.81	39.06	0.00	236.42	17.08	53.89	87.08
	32.98	116.11	89.63	39.91	26.60	14.92	24.82	26.39	168.04	86.53	0.00	18.49
	7.31	0.98	19.04	3.70	0.00	2.72						

Me	m	F	0.00	29.20	6.86	61.61	106.17	6.81	186.24	32.63	54.16	67.78
	31.58	79.48	51.37	41.77	20.93	10.28	24.54	28.79	143.14	71.57	1.78	21.21
	6.81	0.00	21.16	3.66	0.00	3.38						
Me	m	F	0.00	27.40	38.34	73.90	111.85	24.84	148.46	56.67	100.56	85.06
	10.77	47.89	29.57	16.63	12.94	4.78	6.95	12.81	86.11	45.12	0.00	8.86
	5.51	0.91	14.03	0.00	0.00	0.00						
Me	m	F	0.00	38.34	7.27	36.27	90.75	16.52	109.25	63.21	109.77	62.39
	10.84	39.16	19.10	11.40	9.90	2.62	4.73	7.53	48.15	24.01	0.00	3.92
	2.07	0.00	5.98	1.03	0.00	0.00						
Me	m	F	0.00	24.28	14.44	88.75	253.53	49.05	180.09	157.47	294.53	77.01
	15.06	75.05	30.12	21.33	13.28	5.10	9.53	13.03	74.25	42.22	0.00	8.48
	5.10	0.00	20.47	3.50	0.00	0.00						
Me	m	F	0.00	24.41	5.95	19.15	33.46	6.92	278.82	19.89	38.16	
	156.53	32.35	163.41	146.52	47.44	54.82	41.07	40.47	63.41	421.14	283.20	15.60
	50.99	43.93	13.84	122.47	57.41	0.00	9.55					
Me	m	F	0.00	19.45	18.59	44.85	125.74	16.49	154.81	27.93	98.09	52.81
	17.02	35.22	29.93	18.26	10.39	8.39	9.49	6.86	50.10	40.51	0.00	5.58
	2.53	0.00	6.77	1.91	0.00	0.00						
Co	f	L	0.61	45.07	0.00	0.00	0.00	0.00	10.52	0.00	0.00	
	535.32	23.42	12.30	18.95	59.04	26.83	58.28	173.78	114.28	542.28	789.23	
	108.94	247.15	69.21	55.89	207.22	110.21	31.66	29.37				
Co	f	L	1.16	37.17	0.00	0.00	0.00	0.00	12.66	1.98	1.30	
	503.87	15.37	5.51	18.51	48.53	21.17	66.84	174.82	96.18	477.09	937.31	
	176.75	193.86	68.73	69.31	204.01	152.59	23.15	34.22				
Co	f	L	0.00	22.33	0.00	0.00	0.00	0.00	15.50	0.00	0.00	
	423.31	17.59	7.50	17.32	83.12	24.93	53.93	182.17	172.23	481.50	670.14	
	113.02	235.42	60.63	62.64	190.69	99.29	35.82	25.09				
Co	f	L	1.32	33.26	0.00	0.00	0.00	0.00	7.28	2.87	0.00	
	399.67	13.79	4.30	15.83	37.95	21.62	56.21	137.84	52.56	393.16	727.47	
	134.25	133.98	66.69	59.79	137.18	107.45	37.95	117.71				
Co	f	L	0.00	7.64	0.00	0.00	0.00	0.00	6.21	0.00	0.00	
	560.52	7.00	4.78	13.85	36.32	22.23	49.43	160.52	108.33	630.01	915.62	
	133.12	283.98	74.42	60.23	239.23	146.82	79.65	31.49				
Co	f	L	4.81	6.60	0.00	0.00	0.00	0.00	12.92	4.68	0.00	
	463.37	11.99	6.73	15.86	41.00	19.52	50.00	146.88	99.96	453.57	762.34	
	148.22	241.76	64.75	63.90	201.16	147.77	51.65	48.40				
Co	f	L	2.16	5.20	0.00	0.00	0.00	0.00	9.67	3.93	0.00	
	352.63	11.49	5.79	13.35	36.48	16.99	40.26	151.69	65.24	367.89	618.75	
	106.82	207.56	49.93	45.26	142.17	105.50	47.47	31.91				
Co	f	L	0.00	8.18	0.00	0.00	0.00	0.00	2.65	0.00	0.00	
	327.30	7.82	0.00	9.26	24.95	12.99	40.65	142.39	67.59	400.42	717.26	
	121.41	199.34	51.47	48.77	167.59	151.65	55.86	27.78				
Co	f	L	0.25	11.65	0.00	0.00	0.00	0.00	10.34	5.58	7.68	
	433.30	11.42	8.31	19.93	46.74	21.76	55.28	131.20	51.57	296.07	650.79	
	162.73	211.80	49.36	53.07	90.79	86.97	37.00	17.57				
Co	f	L	0.00	4.37	0.00	0.00	0.00	0.00	8.67	0.62	0.69	
	367.48	15.12	7.28	15.67	34.95	17.20	52.43	170.11	48.13	338.77	746.39	
	165.53	195.08	50.83	52.64	114.49	111.79	35.64	20.87				
Co	f	F	0.64	10.73	0.00	0.00	0.00	0.00	6.20	0.00	0.00	
	359.75	8.88	4.39	10.94	30.70	16.54	40.71	131.44	96.04	581.96	864.46	
	117.18	220.57	73.58	60.57	331.27	273.62	85.88	51.84				

Co	f	F	0.00	22.64	0.00	0.00	0.00	0.00	10.51	0.00	0.00	
	334.19	11.37	3.53	6.84	31.16	11.96	29.48	133.97	42.07	267.35	467.93	88.32
	152.63	33.24	30.57	96.15	88.32	21.88	37.05					
Co	f	F	0.00	7.94	0.00	0.00	0.00	0.00	6.75	2.41	1.98	
	230.37	2.83	1.42	2.46	17.81	5.24	14.41	58.24	33.54	187.53	339.96	55.13
	120.97	24.52	22.39	104.16	96.50	23.48	14.27					
Co	f	F	0.00	9.51	0.00	0.00	0.00	0.00	4.32	5.37	7.38	
	379.04	7.74	3.17	9.29	16.64	20.67	52.18	134.10	129.31	677.28	1140.66	
	175.98	174.47	76.41	67.99	315.45	228.88	56.61	44.15				
Co	f	F	0.00	13.87	0.00	0.00	0.00	0.00	6.04	0.00	0.00	
	201.76	3.29	1.98	3.20	12.30	5.54	13.38	50.59	23.24	133.06	243.87	48.24
	85.90	18.42	16.67	52.25	43.87	11.80	7.66					
Co	f	F	0.00	37.83	0.00	0.00	0.00	0.00	4.57	1.54	3.09	
	307.94	7.60	8.69	8.40	12.46	13.31	27.89	71.94	46.97	303.09	524.29	92.51
	105.14	45.89	42.80	156.51	151.37	27.71	21.66					
Co	f	F	0.00	5.79	0.00	0.00	0.00	0.00	5.13	0.56	0.75	
	283.91	7.39	5.04	8.14	24.75	12.24	27.44	114.35	72.56	361.18	557.27	87.34
	155.11	42.49	32.94	158.26	113.65	52.99	26.64					
Co	f	F	0.00	6.28	0.00	0.00	0.00	0.00	3.96	1.24	1.34	
	231.31	4.85	1.24	3.12	19.98	6.82	18.05	70.13	39.86	228.78	377.25	65.23
	107.76	26.81	23.89	108.36	94.71	27.25	16.57					
Co	f	F	0.00	8.94	0.00	0.00	0.00	0.00	0.00	2.36	3.85	
	198.72	2.67	1.70	3.08	10.64	6.78	15.26	53.03	32.32	220.30	335.25	47.28
	89.00	23.07	19.32	91.01	77.49	22.61	12.38					
Co	f	F	0.00	3.60	0.00	0.00	0.00	0.00	5.81	3.18	1.66	
	239.83	5.95	0.93	2.90	15.91	4.01	12.59	72.34	25.17	112.59	251.18	48.55
	97.37	13.55	14.94	43.15	45.37	16.18	5.39					
Co	m	L	0.00	166.46	0.00	0.00	0.00	0.00	8.23	0.00	0.00	
	403.56	7.99	0.00	9.65	46.20	0.00	50.55	112.90	63.61	342.09	609.97	
	118.28	143.83	51.66	45.17	152.77	104.51	20.09	20.57				
Co	m	L	0.00	188.90	0.00	0.00	0.00	0.00	4.29	0.00	0.00	
	245.67	1.93	0.00	14.47	46.01	16.23	47.60	116.06	48.61	318.00	671.40	
	134.06	137.34	45.50	43.73	131.88	103.53	24.56	16.06				
Co	m	L	0.00	90.63	0.00	0.00	0.00	0.00	3.91	0.00	0.00	
	305.37	6.74	0.00	7.13	40.04	13.38	27.54	111.33	52.34	305.86	494.92	85.94
	171.29	38.28	34.08	128.61	88.18	41.21	17.38					
Co	m	L	0.00	67.78	0.00	0.00	0.00	0.00	8.25	0.00	0.00	
	440.85	8.85	0.00	9.91	49.17	21.86	33.43	124.43	75.42	390.85	549.32	78.90
	204.54	43.65	35.17	143.04	72.77	32.07	22.01					
Co	m	L	0.00	125.41	0.00	0.00	0.00	0.00	14.85	0.00	0.00	
	492.19	13.64	0.00	18.04	45.43	0.00	65.35	173.27	62.16	440.92	735.97	
	104.51	99.56	51.16	37.95	122.11	77.89	0.00	16.17				
Co	m	L	0.00	97.27	0.00	0.00	0.00	0.00	8.08	0.00	0.00	
	402.40	12.01	0.00	17.63	57.86	56.39	56.50	167.69	77.73	508.13	831.88	
	137.23	232.04	62.50	49.89	181.71	109.44	35.70	22.16				
Co	m	L	0.00	55.31	0.00	0.00	0.00	0.00	5.65	0.00	0.00	
	264.42	7.06	0.00	9.31	29.98	20.72	29.93	92.73	34.96	265.10	444.58	77.13
	127.21	44.58	35.43	106.65	63.68	18.68	11.46					
Co	m	L	0.00	335.43	0.00	0.00	0.00	0.00	19.91	0.00	0.00	
	339.82	27.93	0.00	22.56	58.97	26.50	54.81	142.54	66.46	365.71	616.58	
	118.09	134.97	42.17	34.82	118.09	66.01	14.31	17.26				

Co	m	L	0.00	90.65	0.00	0.00	0.00	0.00	10.23	0.00	0.00	
	414.85	12.76	0.00	16.23	46.94	29.06	42.14	151.17	46.30	304.30	605.62	
	114.34	172.71	42.64	38.72	103.79	81.43	24.32	15.67				
Co	m	L	0.00	547.65	0.00	0.00	0.00	0.00	8.94	0.00	0.00	
	309.61	11.94	0.00	15.02	43.41	17.79	41.93	133.08	48.95	323.56	522.62	92.27
	135.90	41.53	33.21	109.25	69.56	23.25	0.00					
Co	m	F	0.00	55.93	0.00	0.00	0.00	0.00	5.01	0.00	0.00	
	249.15	11.02	0.00	12.92	41.53	14.19	33.62	129.31	84.32	359.25	635.24	
	135.38	244.56	71.12	65.68	237.36	199.44	167.80	24.65				
Co	m	F	0.00	14.30	0.00	0.00	0.00	0.00	4.87	0.00	0.00	
	258.31	12.78	0.00	12.94	74.28	33.47	47.76	120.69	72.12	311.02	586.90	
	150.00	116.61	46.49	41.85	115.58	108.63	17.01	63.10				
Co	m	F	0.00	13.55	0.00	0.00	0.00	0.00	2.24	0.00	0.00	
	202.03	7.71	0.00	5.26	38.67	9.69	17.98	107.97	46.69	338.72	617.77	
	110.21	162.01	42.73	37.62	158.42	148.83	60.29	20.32				
Co	m	F	0.00	279.87	0.00	0.00	0.00	0.00	10.16	0.00	0.00	
	297.66	8.74	0.00	9.43	42.31	12.07	26.77	111.82	55.89	259.01	413.63	68.59
	139.52	32.58	25.60	93.45	70.25	29.21	13.19					
Co	m	F	0.00	10.04	0.00	0.00	0.00	0.00	4.27	0.00	0.00	
	241.76	6.38	0.00	4.44	28.32	7.81	17.85	77.85	53.19	252.01	345.84	46.70
	121.94	26.02	17.56	104.80	62.97	29.57	13.05					
Co	m	F	0.00	13.21	0.00	0.00	0.00	0.00	4.23	0.00	0.00	
	235.61	4.83	0.00	6.37	21.86	12.54	27.23	76.93	98.19	416.77	576.32	87.32
	116.23	49.56	39.24	257.14	182.09	40.71	48.42					
Co	m	F	0.00	66.92	0.00	0.00	0.00	0.00	5.99	0.00	0.00	
	282.87	9.52	0.00	9.78	40.14	0.00	32.99	85.59	56.21	255.14	543.75	
	138.22	106.42	37.41	38.27	136.90	154.63	33.84	35.03				
Co	m	F	0.00	12.77	0.00	0.00	0.00	0.00	5.97	0.00	0.00	
	239.00	6.98	0.00	5.24	42.40	8.08	26.69	88.98	33.81	241.80	528.34	
	113.83	144.56	31.47	32.06	109.14	105.51	30.09	13.05				
Co	m	F	0.00	20.41	0.00	0.00	0.00	0.00	6.90	0.00	0.00	
	204.55	6.90	0.00	9.17	43.79	14.21	29.31	96.76	64.00	290.62	416.34	63.45
	117.52	25.59	18.83	99.59	81.72	33.72	63.45					
Co	m	F	0.00	2.12	0.00	0.00	0.00	0.00	6.31	0.00	0.00	
	253.39	4.72	0.00	4.37	24.37	5.43	17.88	83.48	33.33	190.86	422.18	96.93
	123.72	21.36	26.90	97.70	102.12	24.42	12.63					
Py	f	L	0.00	4.13	0.00	0.00	0.00	0.00	11.79	0.00	0.00	
	370.73	5.01	3.73	12.08	39.69	17.68	43.42	124.36	73.28	463.56	737.03	86.15
	146.17	51.28	40.08	183.20	94.11	18.07	26.42					
Py	f	L	0.92	1.71	0.00	0.00	0.00	0.00	12.67	3.67	4.71	
	414.25	8.17	8.54	18.04	41.17	23.29	52.50	142.71	65.08	464.50	762.58	
	118.75	171.13	70.17	65.92	193.58	117.79	19.33	23.96				
Py	f	L	1.10	8.41	0.00	0.00	0.00	0.00	11.08	0.55	1.24	
	523.62	0.00	6.30	19.72	70.40	25.32	58.59	167.37	107.67	723.44	1043.24	
	119.76	242.88	84.28	59.79	326.56	138.69	38.60	52.02				
Py	f	L	0.00	5.15	0.00	0.00	0.00	0.00	8.53	2.24	2.07	
	504.40	0.00	5.23	9.59	34.74	13.54	28.41	100.09	121.68	415.26	436.41	44.42
	120.45	43.40	12.97	203.34	59.50	20.84	29.29					
Py	f	L	0.55	6.33	0.00	0.00	0.00	0.00	11.13	2.73	1.64	
	503.90	4.72	5.47	12.85	34.17	19.09	47.91	97.77	79.54	438.54	660.13	77.82
	85.63	55.10	39.40	226.24	93.56	11.64	38.73					

Py	f	L	0.00	6.00	0.00	0.00	0.00	0.00	12.47	1.63	1.47	
	484.87	0.00	6.67	17.36	39.99	21.76	45.63	141.02	114.46	619.66	827.64	87.73
	176.53	66.32	46.98	294.84	118.35	33.88	60.29					
Py	f	L	0.79	4.48	0.00	0.00	0.00	0.00	18.47	2.12	3.82	
	491.37	14.61	7.68	16.19	42.92	25.61	47.99	199.92	78.12	536.36	648.61	68.62
	160.73	41.39	28.23	122.25	56.25	13.53	17.43					
Py	f	L	0.58	3.95	0.00	0.00	0.00	0.00	5.86	0.00	0.00	
	194.60	0.00	2.21	5.28	15.44	7.95	17.18	65.24	40.34	233.43	316.31	33.08
	75.33	22.63	16.31	100.23	39.64	10.16	16.48					
Py	f	L	0.00	7.42	0.00	0.00	0.00	0.00	9.48	0.00	0.00	
	427.73	4.43	4.02	10.10	0.00	13.20	29.38	107.22	75.05	374.95	528.45	62.16
	133.61	43.51	34.02	173.20	71.96	17.94	40.21					
Py	f	L	0.00	8.25	0.00	0.00	0.00	0.00	7.72	0.29	0.41	
	358.07	0.00	5.71	12.07	24.27	14.29	34.58	96.30	53.06	372.98	578.97	70.35
	130.51	43.78	32.94	161.23	73.55	23.94	25.01					
Py	f	F	2.30	6.33	0.00	0.00	0.00	0.00	7.07	1.15	0.00	
	353.80	6.21	5.34	14.10	56.19	27.70	45.50	172.83	121.54	700.04	718.41	49.24
	169.34	60.50	32.31	221.91	81.59	29.22	45.62					
Py	f	F	11.39	15.43	0.00	0.00	0.00	0.00	9.29	4.71	7.06	
	370.49	20.09	9.67	16.35	57.59	23.75	38.00	139.81	100.17	530.22	523.79	55.40
	151.20	50.23	0.00	177.64	64.69	24.97	26.94					
Py	f	F	3.19	2.82	0.00	0.00	0.00	0.00	7.98	0.00	0.00	
	523.45	4.73	5.40	13.26	66.18	26.46	40.33	159.73	144.44	789.44	694.05	64.58
	161.39	77.29	26.09	305.03	99.75	23.27	45.73					
Py	f	F	1.85	1.63	0.00	0.00	0.00	0.00	5.24	0.00	0.00	
	293.00	3.43	3.61	7.97	34.65	16.25	25.28	100.35	81.51	448.26	423.38	36.06
	124.26	41.22	11.71	163.23	55.22	16.56	22.32					
Py	f	F	7.85	13.39	0.00	0.00	0.00	0.00	22.03	8.82	0.00	
	677.64	0.00	15.20	26.56	80.51	41.15	59.31	211.36	198.57	797.60	670.95	
	101.85	252.47	97.23	0.00	269.93	86.70	33.39	31.04				
Py	f	F	18.26	20.93	0.00	0.00	0.00	0.00	16.46	7.16	8.96	
	574.59	17.62	17.87	31.65	80.19	63.39	78.48	288.27	203.31	1374.73		
	996.11	79.26	209.93	112.66	25.07	354.38	110.95	30.43	43.14			
Py	f	F	3.57	9.45	0.00	0.00	0.00	0.00	7.49	1.61	2.90	
	305.65	5.22	3.76	9.45	44.12	15.61	30.67	122.78	90.35	445.45	546.82	51.49
	145.61	41.57	27.06	166.82	61.33	22.12	21.96					
Py	f	F	2.44	14.65	0.00	0.00	0.00	0.00	8.50	3.62	2.40	
	296.38	7.95	7.56	16.06	42.01	25.24	43.35	136.50	104.57	662.52	716.06	65.43
	161.34	60.51	13.70	215.35	77.99	31.02	27.83					
Py	f	F	2.32	11.19	0.00	0.00	0.00	0.00	7.10	0.00	1.41	
	324.10	5.26	4.28	9.46	38.34	16.05	29.71	105.65	97.61	449.76	512.09	54.47
	171.90	48.90	0.00	173.74	67.27	27.83	25.75					
Py	f	F	1.04	5.12	0.00	0.00	0.00	0.00	7.62	0.00	0.00	
	473.74	0.00	2.70	6.33	24.30	13.73	19.93	97.17	84.56	424.30	361.55	28.92
	107.66	38.91	11.69	171.91	56.35	23.26	23.35					
Py	m	L	59.02	43.27	0.00	0.00	0.00	0.00	3.27	0.00	0.00	
	252.68	2.93	2.83	8.49	25.51	13.85	31.32	73.80	58.73	381.95	472.73	46.10
	69.12	41.22	27.51	173.22	65.07	10.83	21.85					
Py	m	L	45.77	59.75	0.00	0.00	0.00	0.00	7.82	0.00	6.15	
	305.06	16.50	8.91	17.77	55.20	20.99	41.12	121.33	83.84	486.14	643.36	68.89
	158.77	50.20	34.96	200.86	75.04	27.72	24.73					

Py	m	L	78.41	68.19	0.00	0.00	0.00	0.00	7.75	0.00	6.26	
	312.73	9.60	7.58	15.45	52.25	16.60	38.24	111.70	77.26	384.88	633.33	75.28
	151.30	48.08	39.18	207.58	95.34	32.55	37.00					
Py	m	L	62.32	87.25	0.00	0.00	0.00	0.00	3.04	0.00	0.00	
	240.25	4.38	2.85	10.09	44.81	15.70	37.87	92.10	66.89	420.27	675.93	77.45
	145.29	46.43	34.71	209.51	97.15	30.54	30.35					
Py	m	L	122.26	120.53	0.00	0.00	0.00	0.00	3.92	0.00	2.12	
	234.09	9.25	5.64	12.70	80.72	24.61	42.95	145.14	123.75	623.67	720.53	68.73
	200.16	65.67	42.87	287.23	101.72	52.66	39.18					
Py	m	L	52.67	101.11	0.00	0.00	0.00	0.00	5.25	0.00	0.38	
	407.76	6.27	5.59	12.87	63.08	21.35	37.83	130.46	103.47	597.16	684.43	61.59
	187.08	57.30	34.80	272.63	95.90	35.81	49.45					
Py	m	L	58.39	74.35	0.00	0.00	0.00	0.00	3.85	0.00	0.69	
	249.22	3.20	2.77	6.27	50.82	12.46	30.75	67.86	61.98	344.68	494.90	51.82
	125.26	36.07	26.30	158.35	62.67	21.54	20.93					
Py	m	L	55.96	119.58	0.00	0.00	0.00	0.00	3.54	0.00	0.00	
	234.71	4.83	3.45	9.05	60.94	14.84	32.79	92.63	63.57	377.50	514.46	49.50
	124.37	41.36	28.20	182.19	73.43	18.14	24.61					
Py	m	L	51.05	63.45	0.00	0.00	0.00	0.00	5.13	0.00	1.35	
	299.81	4.29	3.45	7.74	34.64	12.45	25.31	91.42	64.01	334.50	452.91	43.73
	118.32	34.08	12.77	148.81	60.65	33.10	21.49					
Py	m	L	45.64	59.77	0.00	0.00	0.00	0.00	3.93	0.00	1.23	
	253.75	8.97	4.36	11.86	53.38	15.29	30.47	98.03	61.36	404.24	539.86	56.02
	146.74	40.79	26.90	169.53	66.77	33.97	27.89					
Py	m	F	32.69	41.64	0.00	0.00	0.00	0.00	6.07	0.00	3.76	
	231.16	6.11	6.80	9.41	44.91	15.83	22.36	65.85	79.41	345.68	318.06	28.28
	108.91	31.85	7.41	119.52	33.77	11.29	18.71					
Py	m	F	42.19	35.53	0.00	0.00	0.00	0.00	6.06	0.00	0.00	
	260.06	4.15	3.50	7.14	20.51	8.01	16.79	52.06	41.89	186.07	246.52	28.30
	64.43	18.65	0.00	69.93	34.44	9.52	11.99					
Py	m	F	14.86	17.73	0.00	0.00	0.00	0.00	3.12	0.00	0.00	
	203.34	0.00	2.58	6.14	26.26	11.37	18.20	59.64	60.88	265.31	266.47	22.30
	71.67	22.30	0.00	76.53	23.65	6.61	8.57					
Py	m	F	108.90	90.95	0.00	0.00	0.00	0.00	3.80	0.00	0.00	
	262.48	2.30	4.65	8.35	50.18	16.61	22.76	69.58	73.74	459.28	391.00	30.27
	142.07	39.72	19.81	169.38	51.63	29.46	20.06					
Py	m	F	99.91	27.47	0.00	0.00	0.00	0.00	2.16	0.00	0.38	
	189.71	1.46	2.03	6.23	16.21	9.85	20.63	66.61	53.75	319.66	388.78	38.25
	86.70	34.19	0.00	151.41	63.96	22.66	28.18					
Py	m	F	84.93	34.70	0.00	0.00	0.00	0.00	2.51	0.00	0.26	
	213.38	2.15	2.33	6.03	21.74	10.68	17.85	58.22	59.45	291.78	305.66	24.34
	90.68	26.89	8.36	113.79	39.13	16.76	15.89					
Py	m	F	150.71	91.42	0.00	0.00	0.00	0.00	2.37	0.00	0.00	
	176.84	8.06	3.71	10.26	37.82	17.90	30.83	83.74	67.31	461.36	515.91	44.16
	106.90	39.84	22.98	155.45	56.79	22.55	16.77					
Py	m	F	30.06	28.64	0.00	0.00	0.00	0.00	3.30	0.00	0.00	
	240.85	4.43	4.66	7.78	42.10	14.32	23.52	81.82	88.13	331.14	353.75	38.24
	124.20	35.57	10.80	111.76	34.89	12.84	12.10					
Py	m	F	73.96	59.13	0.00	0.00	0.00	0.00	3.11	0.00	0.34	
	203.91	3.53	2.78	7.58	43.36	13.09	24.72	78.15	80.74	389.41	438.56	44.87
	129.80	42.33	16.95	161.11	55.79	23.92	18.17					

Py	m	F	127.33	72.23	0.00	0.00	0.00	0.00	7.58	0.00	3.44	
	315.33	9.38	8.29	16.04	58.21	20.29	32.13	111.89	94.54	494.16	482.11	53.57
	151.01	56.08	20.24	186.03	60.61	22.20	25.15					

```
##########################
#  1 Data reading      #
##########################

macrolophus <- read.table("S2.Macropholus.data.TXT",sep="\t",header=T)
summary(macrolophus)
dim(macrolophus)
head(macrolophus)



########################################
#  2 Principal components analysis    #
########################################

library(MASS)

macrolophus.pca <- princomp(macrolophus[,4:31],cor=T)
summary(macrolophus.pca)
macrolophus.pca

############################################
#  3 Quadratic discriminant analysis     #
############################################

macrolophus$PC1 <- macrolophus.pca$scores[,1]
macrolophus$PC2 <- macrolophus.pca$scores[,2]
macrolophus$PC3 <- macrolophus.pca$scores[,3]
macrolophus.qda <- qda(species~PC1+PC2+PC3,data=macrolophus)
macrolophus.qda

macrolophus.predict <- predict(macrolophus.qda)
macrolophus.predict


macrolophus$QDA.prob.Me <- macrolophus.predict$posterior[,1]
macrolophus$QDA.prob.Co <- macrolophus.predict$posterior[,2]
macrolophus$QDA.prob.Py <- macrolophus.predict$posterior[,3]
macrolophus$QDA.predicted.species <- macrolophus.predict$class

results <-
macrolophus[,c("species","QDA.prob.Me","QDA.prob.Co","QDA.prob.Py","QDA.predicted.species")]
results[,2:4] <- round(results[,2:4],2)
results

correct.prediction <-
table(macrolophus$species,macrolophus$QDA.predicted.species,dnn=c("real","predicted"))
correct.prediction
prop.table(correct.prediction,1)
prop.table(correct.prediction,2)

total.correct <- sum(diag(correct.prediction))/sum(correct.prediction)
```

```
total.correct

total.error <- 1-total.correct
total.erro

#####################################################
#  4 Quadratic discriminant analysis   by sex    #
#####################################################

#4.1  males

macrolophus.male <- macrolophus[macrolophus$sex=="m",]


macrolophus.pca <- princomp(macrolophus.male[,4:31],cor=T)
summary(macrolophus.pca)
macrolophus.pca

macrolophus.male$PC1 <- macrolophus.pca$scores[,1]
macrolophus.male$PC2 <- macrolophus.pca$scores[,2]
macrolophus.male$PC3 <- macrolophus.pca$scores[,3]
macrolophus.qda <- qda(species~PC1+PC2+PC3,data=macrolophus.male)
macrolophus.qda

macrolophus.predict <- predict(macrolophus.qda)
macrolophus.predict


macrolophus.male$QDA.prob.Me <- macrolophus.predict$posterior[,1]
macrolophus.male$QDA.prob.Co <- macrolophus.predict$posterior[,2]
macrolophus.male$QDA.prob.Py <- macrolophus.predict$posterior[,3]
macrolophus.male$QDA.predicted.species <- macrolophus.predict$class

results <-
macrolophus.male[,c("species","QDA.prob.Me","QDA.prob.Co","QDA.prob.Py","QDA.predicted.species")]
results[,2:4] <- round(results[,2:4],2)
results

correct.prediction <-
table(macrolophus.male$species,macrolophus.male$QDA.predicted.species,dnn=c("real","predicted"))
correct.prediction
prop.table(correct.prediction,1)
prop.table(correct.prediction,2)

total.correct <- sum(diag(correct.prediction))/sum(correct.prediction)
total.correct

total.error <- 1-total.correct
total.error

#4.2  females
```

```
macrolophus.female <- macrolophus[macrolophus$sex=="f",]


macrolophus.pca <- princomp(macrolophus.female[,c(4:5,7:8,10:31)],cor=T)
summary(macrolophus.pca)
macrolophus.pca

macrolophus.female$PC1 <- macrolophus.pca$scores[,1]
macrolophus.female$PC2 <- macrolophus.pca$scores[,2]
macrolophus.female$PC3 <- macrolophus.pca$scores[,3]
macrolophus.qda <- qda(species~PC1+PC2+PC3,data=macrolophus.female)
macrolophus.qda

macrolophus.predict <- predict(macrolophus.qda)
macrolophus.predict


macrolophus.female$QDA.prob.Me <- macrolophus.predict$posterior[,1]
macrolophus.female$QDA.prob.Co <- macrolophus.predict$posterior[,2]
macrolophus.female$QDA.prob.Py <- macrolophus.predict$posterior[,3]
macrolophus.female$QDA.predicted.species <- macrolophus.predict$class

results <-
macrolophus.female[,c("species","QDA.prob.Me","QDA.prob.Co","QDA.prob.Py","QDA.predicted.species
")]
results[,2:4] <- round(results[,2:4],2)
results

correct.prediction <-
table(macrolophus.female$species,macrolophus.female$QDA.predicted.species,dnn=c("real","predicte
d"))
correct.prediction
prop.table(correct.prediction,1)
prop.table(correct.prediction,2)

total.correct <- sum(diag(correct.prediction))/sum(correct.prediction)
total.correct

total.error <- 1-total.correct
total.error


##############################################
#  5 Prediction for new individuals      #
##############################################

library(mnormt)
tomatomale <-
c(9.3359225,35.24040853,0,0,0,0,2.620230036,4.143500323,0,227.0959237,6.803160605,
1.675141163,5.174382195,27.78788589,33.50457571,21.56733939,80.30059764,62.56136815,290.
5071812,451.0842287,
```

```
81.21811111,161.8425264,35.04168524,11.66117306,149.6672654,57.47777861,23.24498038,20.6
8211748)
tomatofemale <-
c(0.465840529,33.08057566,0,1.69478331,0,2.434254523,13.35527135,3.219487532,0.805361245,
623.4437673,21.48072989,5.886884465,11.75026023,17.20756204,32.65718085,44.08958108,133.
3946882,100.2624269,668.2018557,
775.8341467,86.4138592,143.781386,65.40976156,43.59786875,237.3767564,101.9367407,24.226
07705,38.82679428)
macrolophus.new <- rbind(tomatomale,tomatofemale)
source("S4.Macrolophus.predict.R")
predict.QDA(macrolophus.pca,macrolophus.new,macrolophus,macrolophus$species)
```

```
predict.QDA <- function(object.pca,new.individuals,data,species)
{
                data <- as.matrix(data[,4:31])
                gravity.center <- apply(data,2,mean)
                centered.data <- data -
matrix(gravity.center,byrow=T,nrow=nrow(data),ncol=ncol(data))
                variances <- apply(data,2,var) * (nrow(data)-1) / nrow(data)
                deviations <- sqrt(variances)
                centered.reduced.data <- centered.data /
matrix(deviations,byrow=T,nrow=nrow(data),ncol=ncol(data))

                X <- centered.reduced.data
                M <- diag(1,nrow=ncol(X))
                D <- diag(1/nrow(X),nrow=nrow(X))

                #svd decomposition
                eigen.X <- eigen(t(X) %*% D %*% X %*% M)

                #projections for individuals
                F.proof <- X %*% M %*% eigen.X$vectors

                #prediction for new individuals

                X.star <- new.individuals
                X.star <- X.star - matrix(gravity.center,byrow=T,nrow=nrow(X.star),ncol=ncol(X.star))
                X.star <- X.star / matrix(deviations,byrow=T,nrow=nrow(X.star),ncol=ncol(X.star))
                F.star <- X.star %*% M %*% eigen.X$vectors  #projections for new individuals

                mu.Me.proof <- apply(F.proof[species=="Me",1:3],2,mean)
                sigma.Me.proof <- var(F.proof[species=="Me",1:3])
                mu.Co.proof <- apply(F.proof[species=="Co",1:3],2,mean)
                sigma.Co.proof <- var(F.proof[species=="Co",1:3])
                mu.Py.proof <- apply(F.proof[species=="Py",1:3],2,mean)
                sigma.Py.proof <- var(F.proof[species=="Py",1:3])

                prob.X.star.cond.Me <-
dmnorm(x=F.star[,1:3],mean=mu.Me.proof,varcov=sigma.Me.proof)
                prob.X.star.cond.Co <-
dmnorm(x=F.star[,1:3],mean=mu.Co.proof,varcov=sigma.Co.proof)
                prob.X.star.cond.Py <-
dmnorm(x=F.star[,1:3],mean=mu.Py.proof,varcov=sigma.Py.proof)

                prob.Me <- prop.table(table(macrolophus$species))["Me"]
                prob.Co <- prop.table(table(macrolophus$species))["Co"]
                prob.Py <- prop.table(table(macrolophus$species))["Py"]

                prob.X.star <- prob.Me * prob.X.star.cond.Me + prob.Co * prob.X.star.cond.Co + prob.Py
* prob.X.star.cond.Py

                prob.Me.cond.X.star <- prob.Me * prob.X.star.cond.Me /prob.X.star
                prob.Co.cond.X.star <- prob.Co * prob.X.star.cond.Co /prob.X.star
                prob.Py.cond.X.star <- prob.Py * prob.X.star.cond.Py /prob.X.star
```

```
        results <- cbind(prob.Me.cond.X.star,prob.Co.cond.X.star,prob.Py.cond.X.star)
        rownames(results) <- rownames(X.star)
        results <- round(results,3)
        results
}
```