Clustering of loci controlling species differences in chemical bouquets of sympatric *Heliconius* butterflies

Kelsey J. R. P. Byers[1,2,*], Kathy Darragh[1,2,a,*], Sylvia Fernanda Garza[2,b], Diana Abondano Almeida[2,c], Ian A. Warren[1], Pasi M. A. Rastas[3], Richard M. Merrill[2,4], Stefan Schulz[5], W. Owen McMillan[2], Chris D. Jiggins[1,2,6]

[1] *Department of Zoology, University of Cambridge, Cambridge, United Kingdom*

[2] *Smithsonian Tropical Research Institute, Panama, Panama*

[3] *Institute of Biotechnology, University of Helsinki, Helsinki, Finland*

[4] *Division of Evolutionary Biology, Ludwig-Maximilians-Universität München, Munich, Germany*

[5] *Institute of Organic Chemistry, Department of Life Sciences, Technische Universität Braunschweig, Braunschweig, Germany*

[6] *To whom correspondence should be addressed: University of Cambridge, Department of Zoology, Downing Street, Cambridge, CB2 3EJ, United Kingdom. c.jiggins@zoo.cam.ac.uk; Tel: (+44)(0)1223 769021; Fax: (+44)(0)1223 336676*

[a] *Present address: Department of Evolution and Ecology, University of California Davis, Davis, California, USA*

[b] *Present address: Department of Collective Behaviour, Max Planck Institute of Animal Behaviour, Konstanz, Germany & Centre for the Advanced Study of Collective Behaviour, University of Konstanz, Konstanz, Germany*

[c] *Present address: Institute for Ecology, Evolution and Diversity, Goethe Universität, Frankfurt, Germany*

** Kelsey J.R.P. Byers and Kathy Darragh should be considered joint first author*

ORCID K.J.R.P.B.: 0000-0002-0923-8264
ORCID K.D.: 0000-0002-9215-014X
ORCID R.M.M.: 0000-0003-4527-9298
ORCID W.O.M.: 0000-0003-2805-2745
ORCID C.D.J.: 0000-0002-7809-062X

Keywords: pheromones; speciation; quantitative trait locus mapping; Lepidoptera; *Heliconius*

Running title: Clustering of pheromone loci in *Heliconius*

Author contributions:

KJRPB and KD should be considered joint first author. CDJ conceived and designed the study. KJRPB and KD performed phenotyping and QTL mapping for wing and genital compounds respectively. KJRPB

performed downstream analyses. KD and RMM designed and performed QTL mapping crosses and KD and IAW prepared sequencing libraries. KD, SFG, DAA, KJRPB, and RMM reared butterflies for mapping experiments. SS provided mass spectrometry libraries and advice on compound identification. PMAR analyzed raw sequencing data, produced phased sequence data, and constructed linkage maps. CDJ and WOM contributed funding and resources and, with RMM, helped design the QTL mapping experiment. KJRPB wrote the manuscript, and all authors contributed to the manuscript editing process. All authors approved the final version of the manuscript.

Acknowledgments:

Permits for research and collection of butterfly stocks were provided by the government of Panama. KJRPB, KD, IAW, and CDJ were funded by the European Research Council (FP7-IDEAS-ERC 339873); KD was additionally funded by a Natural Environment Research Council Doctoral Training Partnership (Grant No. NE/L002507/1) and a Smithsonian Tropical Research Institute Short Term Fellowship; PMAR was funded by the Jane and Aatos Erkko Foundation; WOM was funded by the Smithsonian Tropical Research Institute; and SS was funded by the Deutsche Forschungsgemeinschaft (grant DFG Schu984/13-1).

Data availability:

Data and QTL analysis code have been deposited in Dryad under accession XXX. Sequencing reads leading to linkage map construction have been deposited in the European Nucleotide Archive (ENA) under project PRJEB34160.

Competing interest statement:

The authors declare no competing interests.

Abstract

The degree to which loci promoting reproductive isolation cluster in the genome – *i.e.* the genetic architecture of reproductive isolation - can influence the tempo and mode of speciation. Tight linkage between these loci can facilitate speciation in the face of gene flow. Pheromones play a role in reproductive isolation in many Lepidoptera species, and their role as secondary metabolites decreases the likelihood of pleiotropy associated with many barrier loci. *Heliconius* butterflies use male sex pheromones to both court females (aphrodisiac wing pheromones) and ward off male courtship (male-transferred anti-aphrodisiac genital pheromones), and it is likely that these compounds play a role in reproductive isolation between *Heliconius* species. Using a set of backcross hybrids between *H. melpomene* and *H. cydno*, we investigated the genetic architecture of putative pheromone compound production. We found a set of 40 significant quantitative trait loci (QTL) representing 33 potential pheromone compounds. QTL clustered significantly on two chromosomes, chromosome 8 for genital compounds and chromosome 20 for wing compounds, and chromosome 20 was enriched for potential pheromone biosynthesis genes. There was minimal overlap between pheromone QTL and known QTL for mate choice and color pattern. Nonetheless, we did detect linkage between a QTL for wing androconial area and *optix*, a color pattern locus known to play a role in reproductive isolation in these species. This tight clustering of putative pheromone loci might contribute to coincident reproductive isolating barriers, facilitating speciation despite ongoing gene flow.

Introduction

The genetic architecture of population differences can profoundly affect the evolution and maintenance of new species, especially in the face of ongoing gene flow. For example, large effect loci are expected to facilitate speciation (Via, 2012, Merrill et al., 2019) because they are less likely to be lost to drift (Kimura, 1983) and a single large effect locus can lead to substantial reproductive isolation on its own (e.g. Bradshaw & Schemske, 2003). Similarly, tight physical linkage of barrier loci, *i.e.* those that contribute to reproductive isolation, will promote speciation by impeding the breakdown of genetic associations between traits that isolate emerging species (Felsenstein, 1981, Smadja & Butlin, 2011). As a result, there is considerable interest in finding the genetic basis of traits that contribute to reproductive isolation and understanding both the effect size and distribution of loci across the genome.

Physical linkage between barrier traits has been detected in a variety of taxa, including pea aphids (Hawthorne & Via, 2001), *Laupala* crickets (Wiley et al., 2011), *Heliconius* butterflies (Merrill et al., 2019), *Ficedula* flycatchers (Sæther et al., 2007), stickleback fish (Bay et al., 2017), and *Aquilegia* columbines (Hodges et al., 2002). Barrier traits can include sex pheromones, chemical signals that mediate intraspecific communication important for mating. Due to their critical role in mate attraction, and their ability to convey information about species identity and male quality, pheromones can be important for establishing and maintaining reproductive isolation through relatively simple changes in chemical bouquets (Smadja & Butlin, 2009). As products of secondary metabolism (*i.e.* products not required for survival), alterations in pheromones also have the potential to be both relatively simple at the molecular level and, if such changes are relatively late in the pathway, likely avoid major pleiotropic consequences.

Some of the best studied sex pheromones involved in speciation are those of female Lepidoptera, and in some cases the genetic basis of variation in these pheromones is now well known. In *Ctenopseustis*, for example, differential regulation of a desaturase drives differences in ratios of sex pheromone components between species (Albre et al., 2012). Desaturases are also important in *Ostrinia* (Roelofs et al., 2002, Sakai et al., 2009, Fujii et al., 2015) and *Helicoverpa* (Li et al., 2015, 2017). Fatty acyl-CoA reductases (FARs) are responsible for species-specific pheromone alterations in *Ostrinia nubialis* strains (Lassance et al., 2010, 2013). Within single species, desaturases and FARs have received the most attention, being functionally characterized in a variety of Lepidoptera including *Yponomeuta* (Liénard et al 2010), *Agrotis* (Ding & Löfstedt, 2015), *Manduca* (Buček et al., 2015), *Bombyx* (Moto et al., 2004), and *Antheraea* (Wang et al., 2010). In some cases, variation in these female pheromones have also been

112 shown to contribute to reproductive isolation (e.g. Liebherr & Roelofs, 1975, Wu et al., 1999, Emelianov
113 et al., 2001; reviewed in Smadja & Butlin, 2009).

114 In contrast, male pheromones in Lepidoptera appear more varied, with a diverse array of compound
115 classes (Conner & Iyengar, 2016, Löfstedt et al., 2016). Many male Lepidopteran pheromones resemble
116 plant compounds and are thought to be diet-derived, especially terpenoids and pyrrolizine alkaloids, the
117 latter of which can be derived from either larval or adult feeding (Conner & Iyengar, 2016). In
118 comparison to female pheromones, the genetic basis of variation in male sex pheromones has received
119 less attention. In male *Bicyclus anynana*, a desaturase (*Ban-Δ11*) and two FARs, *Ban-wFAR1* and *Ban-*
120 *wFAR2*) produce pheromone components and precursors (Liénard et al., 2014), and in the moth *Ostrinia*
121 *nubialis*, males use the same Δ11-desaturase and Δ14-desaturase as females to produce pheromones
122 (Lassance & Löfstedt, 2009). However, the role of these genes in mediating differences between species
123 remains unclear.

124 Speciation is typically thought to rely on the accumulation of multiple reproductive barriers, involving
125 divergence in many different traits (Butlin & Smadja, 2017). In order to understand the extent of genetic
126 linkage that underlies species differences it will therefore be useful to study taxa in which multiple
127 different traits can be mapped. *Heliconius* butterflies have been well studied in the context of speciation
128 (Bates, 1861, Merrill et al., 2015, Jiggins, 2017), and in particular we know a great deal about the species
129 complex of *Heliconius melpomene* and its sister lineage *Heliconius cydno/timareta*. Genetic loci
130 underlying wing pattern and mate preference have been mapped in this group (Merrill et al., 2010,
131 Jiggins, 2017, Merrill et al., 2019). The additional role of chemical signaling in reproductive isolation has
132 long been suspected (Jiggins, 2008), but only recently studied in any detail. Chemical profiles of the male
133 wing androconia and genital regions (thought to act as aphrodisiacs and anti-aphrodisiacs, respectively)
134 differ between species (Estrada et al., 2011, Mann et al., 2017, Darragh et al., 2020), and are important
135 for mate choice (Mérot et al., 2015, Darragh et al., 2017, González-Rojas et al., 2020). Genital
136 pheromones play a more complex role in mate choice in *Heliconius*, as they are transferred by males to
137 females during mating and subsequently decrease advances by other males (Gilbert, 1976, Schulz et al.,
138 2007, Estrada, 2009).

139 One challenge of studying the entire chemical profile is determining which (combination of) compounds
140 are behaviorally active. The chemical bouquets of wings and genitals are complex, often consisting of
141 30-70 compounds (Mann et al., 2017, Darragh et al., 2017, 2020, Byers et al., 2020, this study), but it
142 remains unknown how many of these are important for signaling. A single wing compound in *H.*

143 *melpomene*, octadecanal, is known to be biologically active in *H. melpomene* and *H. cydno* (Byers et al.,
144 2020), while the genital compound (*E*)-β-ocimene acts as the main anti-aphrodisiac in *H. melpomene*
145 (Schulz et al., 2007). In addition, one bioactive genital compound in *H. cydno* (hexyl 3-methylbutyrate,
146 Estrada, 2009) is known. It seems likely that additional compounds in the complex bouquet (30-70
147 compounds) produced by each of these species are biologically active, but this remains to be tested.
148 The simple genetic control of pheromone components in other systems and their role as secondary
149 metabolism products suggests that they can be altered relatively simply without major pleiotropic
150 consequences; in combination with their importance in inter- and intraspecific mate choice, this makes
151 them ripe material for adaptation and speciation. We here conduct quantitative trait locus (QTL)
152 analyses for wing and genital compounds that differ between *Heliconius melpomene* L. and *H. cydno*
153 Doubleday (see also Byers et al., 2020, Darragh et al., 2020). We then investigated the patterns of QTL
154 distribution across the genome to test for clustering of loci across chromosomes.

155

156 **Methods**

157 *Heliconius melpomene rosina* and *H. cydno chioneus* (hereafter *H. melpomene* and *H. cydno*) and their
158 interspecific hybrids were reared in outdoor insectaries at the Smithsonian Tropical Research Institute in
159 Gamboa, Panama. Initial stocks were established from wild-caught individuals, and outbred stocks were
160 refreshed with more wild-caught individuals as necessary. Sexes were kept separately and virgin females
161 were kept separate from egg-laying females. To control for host plant effects, all larvae were fed on the
162 same host plant (*Passiflora platyloba* var. *williamsi*), and adults were maintained on a 20% sugar
163 solution and *Gurania eriantha*, *Psiguria triphylla*, *Psiguria warscewiczii*, and *Psychotria poeppigiana* as
164 pollen sources. Pheromones were collected from male individuals (both parental species and offspring
165 of mapping crosses) following (Darragh et al. 2017). Wing and genital data from *H. melpomene* and wing
166 data from *H. cydno* have previously been published in (Darragh et al., 2017, Darragh et al., 2019, Byers
167 et al., 2020). Briefly, males of 7-14 days post eclosion had their wings and genitals removed and soaked
168 in dichloromethane plus 1 ng/µL 2-tetradecyl acetate (DCM+IS) for one hour, after which the DCM+IS
169 was removed to a fresh vial and stored at -20°C before gas chromatography-mass spectrometry analysis.
170 Bodies were collected and preserved in 20% dimethyl sulfoxide (DMSO) and 0.25 M EDTA (pH 8.0) and
171 stored at –20°C for later DNA extraction.

172 Wing and genital extractions were run on an Agilent 7890B gas chromatograph coupled with an Agilent
173 5977 mass spectrometer (GC-MS) with electron ionization (Agilent Technologies, California, USA) with
174 an HP-5MS capillary column (30m length, 0.25mm inner diameter) and helium as the carrier gas (1.2
175 mL/min). Injection used an Agilent ALS 7693 autosampler and was splitless, with an inlet temperature of
176 250°C. The oven was first held isothermally at 50°C for five minutes, then increased at 5°C/minute to
177 320°C and again held isothermally for five minutes. Compounds were identified with a custom MS
178 library and quantified using the internal standard area. Only compounds that could be definitively
179 identified were included in the data set. To be included in the mapping experiment, compounds had to
180 be present in at least two thirds of the individuals from one of the parent species, i.e. they had to be
181 "typical" *H. melpomene* or *H. cydno* compounds. A limited set of wing compounds (three
182 methyloctadecanals and two henicosenes from *H. melpomene* and *H. cydno* respectively) were excluded
183 from the wing analysis due to close retention times in the GC-MS data preventing definitive
184 identification of the species phenotype in the mapping individuals. Absolute androconial area and
185 absolute hindwing area were measured in the GNU Image Manipulation Program (GIMP) as pixel counts,
186 then absolute androconial area was divided by hindwing area to produce the percentage of the
187 hindwing taken by the androconial region, the relative androconial area.

188 Mapping crosses consisted of backcrosses of either an *H. cydno* or *H. melpomene* mother and an F_1
189 father, as F_1 females are normally sterile. Ten families were constructed as backcrosses to *H. melpomene*
190 (representing 89 individuals for the wing phenotype, 81 individuals for the genital phenotype, and 78
191 individuals for the hindwing relative androconial area phenotype) and fifteen families were constructed
192 as backcrosses to *H. cydno* (127 individuals for the wing, 114 individuals for the genital phenotype, and
193 124 individuals for the hindwing relative androconial area phenotype) for a total of 216 and 195 hybrid
194 individuals for wing and genital studies respectively. DNA extraction, library preparation, and QTL map
195 construction are detailed in (Byers et al., 2020, Darragh et al., bioRxiv). Briefly, Qiagen DNeasy kits
196 (Qiagen) were used for DNA extraction, and individuals then genotyped either by RAD-seq or low-
197 coverage whole genome sequencing using nextera-based libraries (Picelli et al., 2014, Davey et al., 2017,
198 Merrill et al., 2019). Samples were sequenced by HiSeq 3000 (Illumina) by BGI (China). Linkage mapping
199 was conducted using the standard Lep-MAP3 (LM3) pipeline (Rastas, 2017). The initial linkage groups
200 and marker order were constructed based on the *H. melpomene* genome for 21 chromosomes. This
201 resulted in a linkage map with 447,818 SNP markers, which was then evenly thinned by a factor of ten to
202 44,782 markers to ease computation.

203 To determine which cross direction to use for each compound, we inspected the distribution of the
204 compound phenotype in the parents, F_1 individuals, and the backcross individuals. Compounds that
205 were present in equal amounts in both parent species (assessed using a Kruskal-Wallis test) were not
206 mapped. Once cross direction was determined, the remaining compounds were log-transformed to
207 approximate normality before being regressed against the linkage map using R/qtl2 (Broman et al.,
208 2018). Due to the family structure present in our crosses, we additionally included a kinship matrix
209 calculated by R/qtl2 using the LOCO (leave one chromosome out) method. Permutation testing was
210 used (with 1000 replicates) to determine QTL threshold. Compounds are likely to be correlated and our
211 data set represents multiple testing of the same mapping crosses, therefore, we used Bonferroni
212 correction separately for each cross direction to establish a second, more conservative, LOD threshold
213 for significance by dividing the threshold by the number of compounds (traits) mapped in the relevant
214 backcross direction. No correction was used for relative androconial area mapping as the trait was
215 mapped alone due to its general lack of correlation with individual volatile compounds. QTL confidence
216 intervals were obtained using bayes_int. To identify clustering of QTL, we first calculated the expected
217 number of QTL per chromosome, taking into account the chromosome's length in cM, then compared
218 that against the observed distribution of QTL per chromosome using a chi-squared test, computing p-
219 values with Monte Carlo permutation with B = 10,000 permutations. Examination of the residuals was
220 used to determine which chromosome(s) displayed clustering, with any residual over 2 considered
221 significant, as in (Erickson et al., 2016). We also repeated these tests taking into account chromosome
222 length in basepairs because the relationship between map length and physical length can differ across
223 chromosomes. The results were unchanged.

224 As our linkage map is based on whole genome sequencing data, we were able to recover the basepair
225 intervals corresponding to the QTL centimorgan (cM) confidence intervals (defined as the outermost
226 markers at those cM positions) as well as their peaks (defined as the first marker in at that cM position).
227 Lepbase (Challis et al., 2016) was queried to identify genes within the main interval on chromosome 20
228 for wing fatty acid-derived compound production (identified as the minimum overlapping window of all
229 Bonferroni-significant compound confidence intervals). Genes were searched against the nr (non-
230 redundant) protein database using BLASTp (Altschul et al., 1990) to obtain putative functional
231 annotations.

232

233 **Results**

234 We first looked at the distribution of compounds in *H. melpomene*, *H. cydno*, and their hybrids. Analysis
235 of the pheromones found a total of 31 compounds in the wings and 68 compounds in the genitals across
236 the two species that were possible to map (i.e. not unknown or overlapping in the gas chromatography-
237 mass spectrometry trace) (Figure 1, Table S1). There was limited overlap in compounds between the
238 wings and genitals, with only the straight-chain alkanes henicosane, docosane, tricosane, pentacosane,
239 and hexacosane (all of which may be part of the normal whole-body cuticular hydrocarbon profile), the
240 unsaturated aldehyde (*Z*)-11-icosenal, and the nitrogenous aromatic benzyl cyanide found in both body
241 regions. A subset of compounds (6 in the wings and 18 in the genitals) were not mapped as they did not
242 differ significantly between the parental species. The rest showed significant differences and thus had
243 adequate variation for QTL mapping to be feasible. Of the wing compounds, 11 showed segregation in
244 backcrosses to *H. cydno* and thus were mapped with those individuals; nine were mapped in
245 backcrosses to *H. melpomene*; and five showed segregation in both backcross directions and thus were
246 mapped separately in backcrosses to each species. For the genital compounds, 19 were mapped in
247 backcrosses to *H. cydno*; 17 in backcrosses to *H. melpomene*; and 14 in both directions.

248 Most compounds mapped did not produce significant QTL at either the $p = 0.05$ or the Bonferroni-
249 adjusted cutoff levels (the latter taking into account the number of compounds mapped) (Figures S1 and
250 S2). A total of 21 genital compounds and 14 wing compounds produced significant QTL, including three
251 peaks already published in (Byers et al., 2020, Darragh et al., bioRxiv) for octadecanal, 1-octadecanol,
252 and (*E*)-β-ocimene. Of these QTL, nine genital and ten wing compounds were significant at the more
253 conservative Bonferroni threshold (Tables 1 and 2; Table S2). The higher drop-out rate for genital
254 compounds after Bonferroni correction is likely the result of the more stringent correction due to a
255 higher number of compounds corrected for. We also found a single significant QTL on chromosome 18
256 for relative hindwing androconial area in backcrosses to *H. cydno* (Figure S3).

257 Compounds that showed significant QTL tended to be of specific chemical classes. For the wings, of the
258 14 wing compounds with significance at $p = 0.05$, 12 were fatty acid-derived compounds (FADs),
259 including alkanes as well as oxygenated alkanes and alkenes (aldehydes and alcohols). The remaining
260 two significant wing compounds were both nitrogenous aromatics. For the genitals, 10 compounds
261 producing significant QTL were macrolides and eight were FADs (two alkanes, four saturated esters, and
262 two unsaturated esters). The remaining three compounds were again nitrogenous aromatics (two
263 compounds) and one terpene. This is in general agreement with the distribution of compound types in
264 both body regions. The mapped wing compounds comprise 18 fatty acid-derived compounds, three

265 nitrogenous aromatics, two non-nitrogenous aromatics, and two terpenoids, so the finding of significant
266 QTL mostly for fatty acid-derived compounds is not surprising. For the genitals, the mapped compounds
267 consist of 22 fatty acid-derived compounds, 21 macrolides, four terpenoids, two nitrogenous aromatics,
268 and one lactone.

269 Significant quantitative trait loci showed some clustering on specific chromosomes (Figure S4). Notably,
270 these clusters did not overlap with known QTL for color pattern, mate choice, or other traits, which
271 generally tend not to be clustered in the genome, apart from a close association between mate choice
272 and a wing patterning gene (Figure 2). When either all QTL significant at $p < 0.05$ or Bonferroni-
273 corrected significant QTL were combined with known QTL, clustering was statistically significant (χ^2 =
274 64.036 for Bonferroni-significant QTL, χ^2 =78.524 for all significant QTL, $p < 0.001$ for both) for
275 chromosomes 8 and 20. In particular, all but one QTL for FAD compounds in the wings mapped to
276 chromosome 20 (the exception mapped to chromosome 1 as well as having a peak on chromosome 20).
277 The other significant wing compounds (all nitrogenous aromatics) mapped to chromosomes 10, 17, 19,
278 and the Z chromosome (chromosome 21). QTL on chromosome 20 were broadly overlapping, in part
279 due to our relatively small mapping populations producing broad Bayesian confidence intervals for QTL
280 location. In addition, the peaks themselves were strongly overlapping on the latter half of chromosome
281 20 (Figure 3). This clustering was statistically significant (χ^2 = 178.39 for all significant QTL, χ^2 = 113.02 for
282 Bonferroni-corrected significant QTL, $p < 0.001$ in both cases) and residual analysis identified
283 chromosome 20 as the sole outlier. When broken down by chemical class, only FAD compounds
284 clustered significantly, again on chromosome 20 (χ^2 = 232.49 for all significant QTL, $p < 0.001$), and this
285 was also significant when only Bonferroni-corrected significant QTL were tested (χ^2 = 137.55, $p < 0.001$).

286 Similar compound class-specific clustering was seen to a lesser extent for the genitals, where six of the
287 ten macrolides had QTL on chromosome 8, with the others more broadly dispersed on chromosomes 2,
288 7, 13, 14, and 20. The fatty acid-derived compounds (mostly esters) mapped more broadly across the
289 genome than in the wings, with QTL on chromosomes 1, 2, 3, 6, 8, 9, 14, and 20. Clustering was again
290 statistically significant (χ^2 = 46.037 for all significant QTL, χ^2 = 101.83 for Bonferroni-corrected significant
291 QTL, $p = 0.0017$ and $p < 0.001$ respectively) and residual analysis identified chromosomes 8 and 14 as
292 the outliers when all QTL were considered, and chromosome 8 when only Bonferroni-corrected QTL
293 were included. The clustering on chromosome 8 is solely due to the presence of macrolide QTL on this
294 chromosome, while chromosome 14 contains QTL for both a macrolide and three FAD compounds.
295 When broken down by chemical class, only macrolide compounds clustered significantly in the genitals,

296 again on chromosome 8 (χ^2 = 58.079 for all significant QTL, $p < 0.001$), and this was also significant when
297 only Bonferroni-corrected significant QTL were tested (χ^2 = 110.88, $p < 0.001$). Again, the QTL peaks
298 themselves were strongly overlapping in the middle of chromosome 8 (Figure 3) rather than being
299 dispersed across the chromosome. Two compounds had significant QTL in both wings and genitals:
300 henicosane and benzyl cyanide. Both data sets showed QTL on the same chromosomes (20 and 17
301 respectively, with additional QTL for benzyl cyanide in the wings on other chromosomes). The
302 henicosane QTL were within 1.5cM of one another, while the benzyl cyanide QTL on chromosome 17
303 were 2.3cM apart.

304 The strong clustering of QTL for wing fatty acid-derived compounds on chromosome 20 suggested a
305 common gene or set of linked genes responsible for fatty acid-derived compound metabolism or its
306 regulation. We searched the region between 43.11 - 56.37 cM (9,398,348 - 14,585,564 basepairs) for
307 potential candidate genes that could be responsible for this clustering. This region spanned the
308 confidence intervals for all Bonferroni-significant QTL apart from 1-octadecanol and contained 363
309 genes. Annotated proteins from this region (from Lepbase, Challis et al., 2016) were searched using
310 BLASTp against the NCBI non-redundant (nr) database, revealing a total of 20 genes potentially involved
311 in the biosynthesis of the wing compounds within this confidence interval on chromosome 20. These
312 genes include two acetyl-CoA carboxylases (involved in primary metabolism), fifteen fatty acyl-CoA
313 reductases (FARs), and two alcohol dehydrogenases (the latter two families involved in both primary
314 and secondary metabolism), none of which have been functionally characterized. Comparing our
315 candidates to those identified in (Liénard et al., 2014), none of our FARs fall within the pheromone gland
316 FAR (pgFAR) clade, whose members fall on chromosome 19. Inspecting the genome for additional
317 secondary metabolism fatty acid-derived metabolism genes, we identified clusters of FARs on
318 chromosomes 19 and 20 (Figure 4), with a more even distribution of alcohol dehydrogenases and
319 desaturases across the genome. Notably, we did not identify any QTL on chromosome 19 for fatty acid-
320 derived compound production in the wings.

321

322 **Discussion**

323 The genetic basis for traits contributing to reproductive isolation remains poorly studied. This is
324 especially the case for differences in chemical signaling between closely related non-model species, in
325 particular for species with complex chemical signatures where multiple components may play a role in

326 mate choice or species recognition. Here, we mapped a total of 40 QTL controlling the segregation of
327 wing and genital pheromone bouquet compounds in crosses between *Heliconius melpomene* and its
328 close relative *H. cydno*. There was significant evidence for non-random clustering of QTL, most notably
329 for wing compounds on chromosome 20, which showed broadly overlapping QTL with closely spaced
330 peaks. There could be a number of reasons for this clustering of loci. First, theory predicts that
331 mutations controlling species differences may accumulate, and persist, in regions already resistant to
332 gene flow, such that recombination is less likely to disrupt species differences (Smadja & Butlin, 2011).
333 Second, functionally similar genes are likely to occur in close linkage due to *e.g.* tandem gene
334 duplication, such that chemically similar compounds could be regulated by linked genes due to
335 functional similarities between their synthesis pathways (Osbourn, 2010). Finally, related compounds
336 could differ between species as a pleiotropic effect of the same mutation if they are downstream
337 products of the same biochemical pathway or in the case of a multifunctional enzyme. The end result of
338 all of these processes is linkage that can facilitate speciation by impeding recombination that will
339 otherwise disrupt the genetic associations between traits that characterize emerging species.

340 In the case of the compounds described here, it is likely that shared biosynthetic origins play an
341 important role in their chromosomal distributions. The majority of the wing compounds found in either
342 species are fatty acid-derived compounds (FADs; 19 of 31), and the majority of these FADs (12 of 19)
343 were associated with a significant QTL on chromosome 20. It is interesting that there was clustering of
344 one class of biosynthetic enzyme in this pathway, the fatty acyl-CoA reductases (FARs), on both
345 chromosomes 19 and 20, but no QTL on chromosome 19 for any wing or genital FAD compound,
346 suggesting the FARs on chromosome 19 are not involved in putative pheromone compound production.
347 Similarly, QTL for genital compounds exhibited slight clustering of loci underlying macrolide compound
348 production on chromosome 8, but otherwise were more evenly spread across the genome, likely due to
349 their greater biosynthetic diversity. The QTL for shared macrolide compounds may also be due to
350 shared biosynthetic origins of these compounds, but the biosynthesis of macrolides is less well
351 understood and consequently we were unable to identify candidate genes on this chromosome. We also
352 found a QTL on chromosome 3 for hexyl 3-methylbutyrate, a known bioactive component of the genital
353 pheromone of *H. cydno* (Estrada, 2009) whose genetic basis was previously unknown. Of note, the QTL
354 for relative androconial area roughly peaks on the gene *optix* (chromosome 18), a locus controlling red
355 color patterns in *Heliconius* (Reed et al., 2011, Van Belleghem et al., 2017). This locus is also known to
356 control male-specific scales in the basal heliconiine *Dryas iuila* and has an expression domain in the

357 hindwing androconia region (Martin et al., 2014). This overlap raises the possibility for coupled control
358 of color pattern and androconial pheromone production differences between *Heliconius* species.

359 Overall, we found no evidence for linkage of any of the chemical QTL clusters with other loci or QTL that
360 have been shown to play a difference in the strong assortative mating observed between *H. melpomene*
361 and *H. cydno*. The major wing patterning genes are located on chromosomes 1 (*aristalless*), 10 (*WntA*),
362 15 (*Cortex*), and 18 (*optix*), only chromosome 1 of which harbors a chemical QTL, but the peak does not
363 include the *aristalless* gene. In addition, major QTL for male mate choice map to chromosome 1, 17, and
364 18 (Merrill et al., 2019), but, similarly, the chemical QTL we identified on chromosomes 1 and 17 do not
365 overlap with the major QTL identified for male mate choice. We might expect an overlap between
366 chemical QTL and female mate choice QTL (as females likely assess males partially based on their wing
367 androconial chemistry), but the latter has not been mapped in this species pair. Nonetheless, the
368 clustering of wing compounds in particular might contribute to pheromone-based speciation,
369 particularly in the face of gene flow (Via, 2012). If more than one pheromone component that maps to
370 chromosome 20 is involved in reproductive isolation, and if genes responsible for their production are
371 tightly linked (or if pleiotropy occurs), this may also facilitate speciation (Felsenstein, 1981, Smadja &
372 Butlin, 2011, Merrill et al., 2019) and prevent the breakup of potentially adaptive pheromone bouquet
373 mixtures. Evidence suggests that mixture effects are important in pheromone processing (Riffell et al.
374 ,2009, Clifford & Riffell, 2013, Lei et al., 2013), and these mixtures might be maintained through tight
375 genetic linkage between biosynthetic loci, facilitated by gene evolutionary events such as tandem
376 duplication. In contrast to the tight clustering seen on chromosome 20 for wing FAD compounds, we
377 saw more limited clustering for the more chemically diverse genital compounds, likely due to their more
378 diverse potential biosynthesis pathways. This might hamper any potential role in reproductive isolation.
379 Given their known function as anti-aphrodisiacs, genital compounds are not so likely to play a major role
380 in isolation between *Heliconius* species (Gilbert, 1976, Schulz et al., 2007, Estrada, 2009).

381 Some of the compounds found on wings and in genitalia might have a plant origin. Several compounds
382 in the wings and genitals did not differ between parental species, and many of these have been shown
383 to be affected by larval diet in *Heliconius melpomene*. Most of the invariant wing compounds are
384 aromatics, which are expected to be plant derived in *Heliconius* (Darragh et al., 2019). Notably, we
385 reared all mapping butterflies on the same larval and adult diets, thus effectively creating a dietary
386 common garden experiment, and so we do not expect larval or adult diet to affect the result of the QTL
387 mapping. Many male moth pheromone components are derived from host plant volatiles with minimal

388 alterations, potentially due to females' pre-existing biases for detecting these compounds during the
389 host plant search (Conner & Iyengar, 2016). However, many of our mapped *Heliconius* compounds do
390 appear to have a genetic basis (i.e. a QTL was found), in line with the biosynthesis of FAD and related
391 compounds by female moths. Breaking our compounds down into four classes (FADs, macrolides,
392 aromatics, and terpenoids), we failed to find QTL (at the $p = 0.05$ level) for 54%, 58%, 43%, and 83% of
393 these classes respectively across both wings and genitals. In particular, terpenoids that differ between
394 species may be the result of differential larval host plant use. It is of course possible that plant derived
395 compounds could show a genetic basis if there were species differences in the genes regulating their
396 uptake, and one example might be icosanal, which has been shown to be influenced by larval host plant
397 (Darragh et al., 2019) but which is also associated with a QTL in this study.

398 This study provides a first step towards better characterizing the genes causing chemical differences
399 between species. Fatty acid-derived compound biosynthesis candidate genes and QTL for FAD
400 compounds are both clustered on chromosome 20. These FAD compounds are produced from acetyl-
401 CoA followed by rounds of chain elongation, desaturation, fatty acyl-CoA reduction, oxidization, and
402 acetyltransferase activity to produce the final pheromone compounds. The fatty acyl-CoA reduction
403 step, key to pheromone biosynthesis, uses FARs to remove the CoA moiety and return an alcohol. More
404 than half of these FAR candidates are found in tandem duplications, as are the two alcohol
405 dehydrogenase candidates. These duplications could have provided raw material for the origin of new
406 biosynthetic function and thus increased pheromone diversity. FARs also cluster on chromosome 19,
407 and these FARs (unlike those on chromosome 20) fall within the pheromone gland FAR (pgFAR)
408 subfamily (Liénard et al., 2014, Löfstedt et al., 2016). However, these phylogenetic analyses of FARs in
409 *Heliconius*, *Bombyx, Danaus,* and *Bicyclus* do not generally show lineage-specific clustering of the
410 chromosome 20 FARs; instead, they are dispersed throughout the tree. This suggests an ancient
411 evolutionary origin of the *Heliconius* chromosome 20 FAR cluster, in contrast to *Heliconius*-specific
412 duplications of chromosome 19 FARs in the pgFAR subfamily which are clustered in the tree.

413 Single FAR and alcohol dehydrogenase genes are missing from *H. cydno* (Byers et al., 2020), suggesting
414 an obvious mechanism for the striking difference in FAD compounds between the two species; these
415 differences may also be due to e.g. differential regulation of FARs or other biosynthetic genes, or
416 mutations within these genes. Alternately, a single FAR gene or small set of genes may be responsible
417 for the biosynthesis of the majority of FAD compounds in *Heliconius*, similar to the situation found with
418 *eloF* in *Drosophila*, where knockdown of the gene produces widely pleiotropic effects on cuticular

419 hydrocarbon profiles (Combs et al., 2018). Similar effects are seen in *Spodoptera litura*, where RNAi of
420 *FAR17* alone alters levels of the four major sex pheromone components (Lin et al., 2017). Another
421 option, though one we consider unlikely, is that this tight cluster of FARs on chromosome 20 serves as a
422 supergene, allowing multiple phenotypic traits (here individual chemical compounds) to co-segregate
423 and maintaining strong integration of the chemical phenotype. Supergenes have been described for
424 other phenotypic morphs in *Heliconius* (Joron et al., 2011), but we are unaware of mechanisms on
425 chromosome 20 that might contribute to supergene formation and maintenance, e.g. a chromosomal
426 inversion.

427 Information on potential biosynthetic gene clustering in other Lepidoptera species is largely lacking, as
428 most studies have produced transcriptomes of pheromone glands rather than linkage maps or
429 transcriptomes mapped to genomes. Additionally, as much work in Lepidoptera has focused on single
430 pheromone components and thus single biosynthetic genes (Groot et al., 2016), clustering of QTL and
431 candidate genes is rarely discussed. The lack of annotated genomes in many lepidopteran species also
432 can prevent in-depth assessment of clustering of candidate genes, requiring mapping QTL linkage groups
433 back to e.g. the *Bombyx* genome. However, such clustering has been seen in *Nasonia* wasps, where both
434 QTL for multiple cuticular hydrocarbon components and candidate desaturase genes overlap strongly
435 across the genome (Niehuis et al., 2011, 2013). Clusters of acetyltransferases and desaturases exist in
436 *Heliothis* (inferred from synteny with *Bombyx mori*), but the acetyltransferases do not appear to
437 underlie the QTL for sex pheromone differences between *H. virescens* and *H. subflexa* (Groot et al.,
438 2013, 2014). Quantitative trait loci for multiple compounds were investigated in *Heliothis*, finding
439 multiple chromosomes responsible for the production of the same compound classes (all FADs), with
440 more limited clustering (Sheck et al., 2006). This is similar in pattern to what we found in the genital
441 pheromone components, although the latter cover multiple compound classes.

442 Understanding the genetic architecture of complex traits can allow us to understand their potential role
443 in speciation. Determining whether the observed clustering QTL on e.g. chromosomes 8 and 20 reflect
444 tight linkage of multiple biosynthetic and/or regulatory genes or a single "master" locus responsible for
445 the biosynthesis or regulation of multiple pheromone components will contribute to our understanding
446 of both pheromone biosynthesis and pleiotropy and linkage in general. With genes in hand, CRISPR (a
447 technique that works in *Heliconius*, Livraghi et al., 2018) could be used to alter gene function, allowing
448 confirmation of the role of pheromone components via behavioral assays. As major effect loci (as these
449 QTL appear to be) and tight linkage (as potentially seen here) can both facilitate the speciation process,

450 especially in the face of gene flow (Felsenstein, 1981, Merrill et al., 2010, Smadja & Butlin, 2011, Via,
451 2012), the QTL identified here could be barrier or speciation genes in *Heliconius*. These specific QTL may
452 be involved in interspecific mate choice and reproductive isolation, and thus the genes underlying them
453 could contribute to the maintenance of species boundaries. We have here identified a total of 40 QTL
454 underlying species differences in 33 potential pheromone components, demonstrated clustering of
455 these QTL for wing and genital compounds on specific chromosomes not linked to known loci for species
456 differences, and identified candidate genes underlying the production of the major chemical class of
457 wing compounds. Together, these findings further our knowledge of chemical ecology, pheromone
458 genetics, and their potential roles in *Heliconius* butterfly mate choice and speciation.

459

460 **References**

461 Albre, J., Liénard, M. A., Sirey, T. M., Schmidt, S., Tooman, L. K., Carraher, C., ...Newcomb, R. D. (2012)
462 Sex Pheromone Evolution Is Associated with Differential Regulation of the Same Desaturase Gene in
463 Two Genera of Leafroller Moths. PLOS Genetics 8, e1002489. http://doi: 10.1371/journal.pgen.1002489

464 Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990) Basic local alignment search tool.
465 Journal of Molecular Biology, 215, 403-410. http://doi: 10.1016/S0022-2836(05)80360-2

466 Bates, H. W. (1862) XXXII. Contributions to an Insect Fauna of the Amazon Valley. Lepidoptera:
467 Heliconidæ. Transactions of the Linnean Society of London, 23, 495-566. http://doi: 10.1111/j.1096-
468 3642.1860.tb00146.x

469 Bay, R. A., Arnegard, M. E., Conte, G. L., Best, J., Bedford, N. L., McCann, S. R., ...Peichel, C. L. (2017)
470 Genetic Coupling of Female Mate Choice with Polygenic Ecological Divergence Facilitates Stickleback
471 Speciation. Current Biology, 27, 3344-3349.e4. http://doi: 10.1016/j.cub.2017.09.037

472 Bradshaw, H. D., & Schemske, D. W. (2003) Allele substitution at a flower colour locus produces a
473 pollinator shift in monkeyflowers. Nature, 426, 176-178. http://doi: 10.1038/nature02106

474 Broman, K. W., Gatti, D. M., Simecek, P., Furlotte, N. A., Prins, P., Sen, Ś., ...Churchill, G. A. (2018) R/qtl2:
475 software for mapping quantitative trait loci with high-dimensional data and multi-parent populations.
476 Genetics, 211, 495-502. http://doi: 10.1534/genetics.118.301595

477 Buček, A., Matoušková, P., Vogel, H., Šebesta, P., Jahn, U., Weißflog., J., ...Pichová, I. (2015) Evolution of
478 moth sex pheromone composition by a single amino acid substitution in a fatty acid desaturase.
479 Proceedings of the National Academy of Sciences USA, 112, 12586-12591. http://doi:
480 10.1073/pnas.1514566112

481 Butlin, R. K., & Smadja, C. M. (2017) Coupling, reinforcement, and speciation. The American Naturalist,
482 191, 155-172. http://doi: 10.1086/695136

483 Byers, K. J. R. P., Darragh, K., Musgrove, J., Abondano Almeida, D., Fernanda Garza, S., Warren, I. A.,
484 ...Jiggins, C. D. (2020) A major locus controls a biologically active pheromone component in *Heliconius*
485 *melpomene*. Evolution, 74, 349-364. http://doi: 10.1111/evo.13922

486 Challis, R. J., Kumar, S., Dasmahapatra, K. K., Jiggins, C. D., & Blaxter, M. (2016) Lepbase: the
487 Lepidopteran genome database. bioRxiv, 056994v1. http://doi: 10.1101/056994

488 Clifford, M. R., & Riffell, J. A. (2013) Mixture and odorant processing in the olfactory systems of insects: a
489 comparative perspective. Journal of Comparative Physiology A, 199, 911-928. http://doi:
490 10.1007/s00359-013-0818-6

491 Combs, P. A., Krupp, J. J., Khosla, N. M., Bua, D., Petrov, D. A., Levine, J. D., & Fraser, H. B. (2018) Tissue-
492 Specific cis-Regulatory Divergence Implicates *eloF* in Inhibiting Interspecies Mating in *Drosophila*.
493 Current Biology, 28, 3969-3975.e3. http://doi: 10.1016/j.cub.2018.10.036

494 Connor, W. E., & V. K. Iyengar (2016) Male pheromones in moths: reproductive isolation, sexy sons, and
495 good genes. In Pheromone Communication in Moths, J. D. Allison and R. Cardé, ed. Oakland, CA, USA:
496 University of California Press.

497 Darragh, K., Byers, K. J. R. P., Merrill, R. M., McMillan, W. O., Schulz, S., & Jiggins, C. D. (2019a) Male
498 pheromone composition depends on larval but not adult diet in *Heliconius melpomene*. Ecological
499 Entomology, 44, 397-405. http://doi: 10.1111/een.12716

500 Darragh, K., Montejo-Kovacevich, G., Kozak, K. M., Morrison, C. M., Figueiredo, C. M. E., Ready, J. S.,
501 ...Jiggins, C. D. (2020) Species specificity and intraspecific variation in the chemical profiles of *Heliconius*
502 butterflies across a large geographic range. Ecology and Evolution. http://doi: 10.1002/ece3.6079

503 Darragh, K., Orteu, A., Byers, K. J. R. P., Szczerbowski, D., Warren, I. A., Rastas, P., ...Jiggins, C. D. (2019b)
504 A novel terpene synthase produces an anti-aphrodisiac pheromone in the butterfly *Heliconius*
505 *melpomene*. bioRxiv, 779678v1. http://doi: 10.1101/779678

506 Darragh, K., Vanjari, S., Mann, F., González-Rojas, M. F., Morrison, C. R., Salazar, C., ...Jiggins, C. D. (2017)
507 Male sex pheromone components in *Heliconius* butterflies released by the androconia affect female
508 choice. PeerJ, 5, e3953. http://doi: 10.7717/peerj.3953

509 Davey, J. W., Barker S. L., Rastas, P. M., Pinharanda, A., Martin, S. H., Durbin, R., ...Jiggins, C. D. (2017) No
510 evidence for maintenance of a sympatric *Heliconius* species barrier by chromosomal inversions.
511 Evolution Letters, 1, 138-154. http://doi: 10.1002/evl3.12

512 Ding, B.-J., & Löfstedt, C. (2015) Analysis of the *Agrotis segetum* pheromone gland transcriptome in the
513 light of sex pheromone biosynthesis. BMC Genomics, 16, 711. http://doi: 10.1186/s12864-015-1909-2

514 Emelianov, I., Drès, M., Baltensweiler, W., & Mallet, J. (2001) Host-induced assortative mating in host
515 races of the larch budmoth. Evolution, 55, 2002-2010. http://doi: 10.1111/j.0014-3820.2001.tb01317.x

516 Erickson, P. A., Glazer, A. M., Killingbeck, E. E., Agoglia, R. M., Baek, J., Carsanaro, S. M., ...Miller, C. T.
517 (2016) Partially repeatable genetic basis of benthic adaptation in threespine sticklebacks. Evolution, 70,
518 887-902. http://doi: 10.1111/evo.12897

519 Estrada, C. (2009) Sexual Behavior, Intraspecific Signaling and the Evolution of Mimicry among Closely
520 Related Species. PhD thesis, University of Texas Austin.

521 Estrada, C., Schulz, S., Yildizhan, S., & Gilbert, L. E. (2011) Sexual selection drives the evolution of
522 antiaphrodisiac pheromones in butterflies. Evolution, 65, 2843-2854. http://doi: 10.1111/j.1558-
523 5646.2011.01352.x

524 Felsenstein, J. (1981) Skepticism towards Santa Rosalia, or why are there so few kinds of animals?
525 Evolution, 35, 124-138. http://doi: 10.1111/j.1558-5646.1981.tb04864.x

526 Fujii, T., Yasukochi, Y., Rong, Y., Matsuo, T., & Ishikawa, Y. (2015) Multiple Δ11-desaturase genes
527 selectively used for sex pheromone biosynthesis are conserved in *Ostrinia* moth genomes. Insect
528 Biochemistry and Molecular Biology, 61, 62-68. http://doi: 10.1016/j.ibmb.2015.04.007

Gilbert, L. E. (1976) Postmating female odor in *Heliconius* butterflies: a male-contributed antiaphrodisiac? Science, 193, 419-420. http://doi: 10.1126/science.935877

González-Rojas, M. F., Darragh, K., Robles, J., Linares, M., Schulz, S., McMillan, W. O., …Salazar, C. (2020) Chemical signals act as the main reproductive barrier between sister and mimetic *Heliconius* butterflies. Proceedings of the Royal Society B: Biological Sciences, 287, 20200587. http://doi: 10.1098/rspb.2020.0587

Groot, A. T., Dekker, T., & Heckel, D. G. (2016) The genetic basis of pheromone evolution in moths. Annual Review of Entomology, 61, 99-117. http://doi: 10.1146/annurev-ento-010715-023638

Groot, A. T., Schöfl, G., Inglis, O., Donnerhacke, S., Classen, A., Schmalz, A., ...Heckel, D. G. (2014) Within-population variability in a moth sex pheromone blend: genetic basis and behavioral consequences. Proceedings of the Royal Society B: Biological Sciences, 281, 20133054. http://doi: 10.1098/rspb.2013.3054

Groot, A. T., Staudacher, H., Barthel, A., Inglis, O., Schöfl, G., Santangelo, R. G., ...Gould, F. (2013) One quantitative trait locus for intra- and interspecific variation in a sex pheromone. Molecular Ecology, 22, 1065-1080. http://doi: 10.1111/mec.12171

Hawthorne, D. J., & Via, S. (2001) Genetic linkage of ecological specialization and reproductive isolation in pea aphids. Nature, 412, 904-907. http://doi: 10.1038/35091062

Hodges, S. A., Whittall, J. B., Fulton, M., & Yang, J. Y. (2002) Genetics of Floral Traits Influencing Reproductive Isolation between *Aquilegia formosa* and *Aquilegia pubescens*. American Naturalist, 159, S51-S60. http://doi: 10.1086/338372

Jiggins, C. D. (2008) Ecological speciation in mimetic butterflies. BioScience, 58, 541-548. http://doi: 10.1641/B580610

Jiggins, C. D. (2017) The ecology and evolution of *Heliconius* butterflies. Oxford, UK: Oxford University Press.

Joron, M., Frezal, L., Jones, R. T., Chamberlain, N. L., Lee, S. F., Haag, C. R., ...ffrench-Constant, R. H. (2011) Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature, 477, 203-206. http://doi: 10.1038/nature10341

556 Kimura, M. (1983) The neutral theory of molecular evolution. Cambridge, UK: Cambridge University
557 Press.

558 Lassance, J.-M., & Löfstedt, C. (2009) Concerted evolution of male and female display traits in the
559 European corn borer, *Ostrinia nubialis*. BMC Biology, 7, 10. http://doi: 10.1186/1741-7007-7-10

560 Lassance, J.-M., Groot, A. T., Liénard, M. A., Antony, B., Borgwardt, C., Andersson, F., ...Löfstedt, C.
561 (2010) Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones.
562 Nature, 466, 486-489. http://doi: 10.1038/nature09058

563 Lassance, J.-M., Liénard, M. A., Antony, B., Qian, S., Fujii, T., Tabata, J., ...Löfstedt, C. (2013) Functional
564 consequences of sequence variation in the pheromone biosynthetic gene *pgFAR* for *Ostrinia* moths.
565 Proceedings of the National Academy of Sciences USA, 110, 3967-3972. http://doi:
566 10.1073/pnas.1208706110

567 Lei, H., Chiu, H.-Y., & Hildebrand, J. G. (2013) Responses of protocerebral neurons in *Manduca sexta* to
568 sex-pheromone mixtures. Journal of Comparative Physiology A, 199, 997-1014. http://doi:
569 10.1007/s00359-013-0844-4

570 Li, R.-T., Ning, C., Huang, L.-Q., Dong, J.-F., Li, X., & Wang, C.-Z. (2017) Expressional divergences of two
571 desaturase genes determine the opposite ratios of two sex pheromone components in *Helicoverpa*
572 *armigera* and *Helicoverpa assulta*. Insect Biochemistry and Molecular Biology, 90, 90-100. http://doi:
573 10.1016/j.ibmb.2017.09.016

574 Li, Z.-Q., Zhang, S., Luo, J.-Y., Wang, C.-Y., Lv, L.-M., Dong, S.-L., & Cui, J.-J. (2015) Transcriptome
575 comparison of the sex pheromone glands from two sibling *Helicoverpa* species with opposite sex
576 pheromone components. Scientific Reports, 5, 9324. http://doi: 10.1038/srep09324

577 Liebherr, J., & Roelofs, W. (1975) Laboratory Hybridization and Mating Period Studies Using Two
578 Pheromone Strains of *Ostrinia nubilalis*. Annals of the Entomological Society of America, 68, 305-309.
579 http://doi: 10.1093/aesa/68.2.305

580 Lin, X., Wang, B., & Du, Y. (2017) Key genes of the sex pheromone biosynthesis pathway in female moths
581 are required for pheromone quality and possibly mediate olfactory plasticity in conspecific male moths
582 in *Spodoptera litura*. Insect Molecular Biology, 27, 8-21. http://doi: 10.1111/imb.12335

583 Livraghi, L., Martin, A., Gibbs, M., Braak, N., Arif, S., & Breuker, C. J. (2018) Chapter Three - CRISPR/Cas9
584 as the Key to Unlocking the Secrets of Butterfly Wing Pattern Development and Its Evolution. Advances
585 in Insect Physiology, 54, 85-115. http://doi: 10.1016/bs.aiip.2017.11.001

586 Liénard, M. A., Hagström, Å. K., Lassance, J.-M., & Löfstedt, C. (2010) Evolution of multicomponent
587 pheromone signals in small ermine moths involves a single fatty-acyl reductase gene. Proceedings of the
588 National Academy of Sciences USA, 107, 10955-10960. http://doi: 10.1073/pnas.1000823107

589 Liénard, M. A., Wang, H.-L., Lassance, J.-M., & Löfstedt, C. (2014) Sex pheromone biosynthetic pathways
590 are conserved between moths and the butterfly *Bicyclus anynana*. Nature Communications, 5, 3957.
591 http://doi: 10.1038/ncomms4957

592 Löfstedt, C., Wahlberg, N., & Millar, J. G. (2016) Evolutionary patterns of pheromone diversity in
593 Lepidoptera. In Pheromone Communication in Moths, J. D. Allison and R. Cardé, ed. Oakland, CA, USA:
594 University of California Press.

595 Mann, F., Vanjari, S., Rosser, N., Mann, S., Dasmahapatra, K. K., Corbin, C., ...Schulz, S. (2017) The scent
596 chemistry of *Heliconius* wing androconia. Journal of Chemical Ecology, 43, 843-857. http://doi:
597 10.1007/s10886-017-0867-3

598 Martin, A., McCulloch, K. J., Patel, N. H., Briscoe, A. D., Gilbert, L. E., & Reed, R. D. (2014) Multiple recent
599 co-options of *Optix* associated with novel traits in adaptive butterfly wing radiations. EvoDevo, 5, 7.
600 http://doi: 10.1186/2041-9139-5-7

601 Merrill, R. M., Dasmahapatra, K. K., Davey, J. W., Dell'Aglio, D. D., Hanly, J. J., Huber, B., ...Yu, Q. (2015)
602 The diversification of *Heliconius* butterflies: what have we learned in 150 years? Journal of Evolutionary
603 Biology, 28, 1417-1438. http://doi: 10.1111/jeb.12672

604 Merrill, R. M., Rastas P., Martin S. H., Melo, M. C., Barker, S., Davey, J., ...Jiggins, C. D. (2019) Genetic
605 dissection of assortative mating behavior. PLOS Biology 17:e2005902. http://doi:
606 10.1371/journal.pbio.2005902

607 Merrill, R. M., Van Schooten, B., Scott, J. A., & Jiggins, C. D. (2010) Pervasive genetic associations
608 between traits causing reproductive isolation in *Heliconius* butterflies. Proceedings of the Royal Society
609 of London B: Biological Sciences, 278, 511-518. http://doi: 10.1098/rspb.2010.1493

610 Moto, K., Suzuki, M. G., Hull, J. J., Kurata, R., Takahashi, S., Yamamoto, M., ...Matsumoto, S. (2004)
611 Involvement of a bifunctional fatty-acyl desaturase in the biosynthesis of the silkmoth, *Bombyx mori*, sex
612 pheromone. Proceedings of the National Academy of Sciences USA, 101, 8631-8636. http://doi:
613 10.1073/pnas.0402056101

614 Mérot, C., Frérot, B., Leppik, E., & Joron, M. (2015) Beyond magic traits: Multimodal mating cues in
615 *Heliconius* butterflies. Evolution, 69, 2891-2904. http://doi: 10.1111/evo.12789

616 Niehuis, O., Büllesbach, J., Gibson, J. D., Pothmann, D., Hanner, C., Mutti, N. S., ...Schmitt, T. (2013)
617 Behavioral and genetic analyses of *Nasonia* shed light on the evolution of sex pheromones. Nature, 494,
618 345-348. http://doi: 10.1038/nature11838

619 Niehuis, O., Büllesbach, J., Judson, A. K., Schmitt, T., & Gadau, J. (2011) Genetics of cuticular
620 hydrocarbon differences between males of the parasitoid wasps *Nasonia giraulti* and *Nasonia*
621 *vitripennis*. Heredity, 107, 61-70. http://doi: 10.1038/hdy.2010.157

622 Osbourn, A. (2010) Secondary metabolite gene clusters: evolutionary toolkits for chemical innovation.
623 Trends in Genetics, 26, 449-457. http://doi: 10.1016/j.tig.2010.07.001

624 Picelli S., Björklund, Å. K., Reinius, B., Sagasser, S., Winberg, G., & Sandberg, R. (2014) Tn5 transposase
625 and tagmentation procedures for massively scaled sequencing projects. Genome Research, 24, 2033-
626 2040. http://doi: 10.1101/gr.177881.114

627 Rastas, P. 2017. Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing
628 data. Bioinformatics, 33, 3726-3732. http://doi: 10.1093/bioinformatics/btx494

629 Reed, R. D., Papa, R., Martin, A., Hines, H. M., Counterman, B. A., Pardo-Diaz, C., ...McMillan, W. O.
630 (2011) *optix* Drives the Repeated Convergent Evolution of Butterfly Wing Pattern Mimicry. Science, 333,
631 1137-1141. http://doi: 10.1126/science.1208227

632 Riffell, J. A., Lei, H., Christensen, T. A., & Hildebrand, J. G. (2009) Characterization and Coding of
633 Behaviorally Significant Odor Mixtures. Current Biology, 19, 335-340. http://doi:
634 10.1016/j.cub.2009.01.041

635 Roelofs, W. L., Liu, W., Hao, G., Jiao, H., Rooney, A. P., & Linn, C. E., Jr. (2002) Evolution of moth sex
636 pheromones via ancestral genes. Proceedings of the National Academy of Sciences USA, 99, 13621-
637 13626. http://doi: 10.1073/pnas.152445399

638 Sakai, R., Fukuzawa, M., Nakano, R., Tatsuki, S., & Ishikawa, Y. (2009) Alternative suppression of
639 transcription from two desaturase genes is the key for species-specific sex pheromone biosynthesis in
640 two *Ostrinia* moths. Insect Biochemistry and Molecular Biology, 39, 62-67. http://doi:
641 10.1016/j.ibmb.2008.10.001

642 Schulz, S., Estrada, C., Yildizhan, S., Boppré, M., & Gilbert L. E. (2007) An Antiaphrodisiac in *Heliconius*
643 *melpomene* Butterflies. Journal of Chemical Ecology, 34, 82-93. http://doi: 10.1007/s10886-007-9393-z

644 Sheck, A. L., Groot, A. T., Ward, C. M., Gemeno, C., Wang, J., Brownie, C., ...Gould, F. (2006) Genetics of
645 sex pheromone blend differences between *Heliothis virescens* and *Heliothis subflexa*: a chromosome
646 mapping approach. Journal of Evolutionary Biology, 19, 600-617. http://doi: 10.1111/j.1420-
647 9101.2005.00999.x

648 Smadja, C. M., & Butlin, R. K. (2009) On the scent of speciation: the chemosensory system and its role in
649 premating isolation. Heredity, 102, 77-97. http://doi: 10.1038/hdy.2008.55

650 Smadja, C. M., & Butlin, R. K. (2011) A framework for comparing processes of speciation in the presence
651 of gene flow. Molecular Ecology, 20, 5123-5140. http://doi: 10.1111/j.1365-294X.2011.05350.x

652 Sæther, S. A., Sætre, G.-P., Borge, T., Wiley, C., Svedin, N., Andersson, G., ...Qvarnström, A. (2007) Sex
653 Chromosome-Linked Species Recognition and Evolution of Reproductive Isolation in Flycatchers.
654 Science, 318, 95-97. http://doi: 10.1126/science.1141506

655 Van Belleghem, S. M., Rastas P., Papanicolaou, A., Martin, S. H., Arias, C. F., Supple, M. A., & Papa, R.
656 (2017) Complex modular architecture around a simple toolkit of wing pattern genes. Nature Ecology and
657 Evolution, 1, 0052. http://doi: 10.1038/s41559-016-0052

658 Via, S. (2012) Divergence hitchhiking and the spread of genomic isolation during ecological speciation-
659 with-gene-flow. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 367,
660 451-460. http://doi: 10.1098/rstb.2011.0260

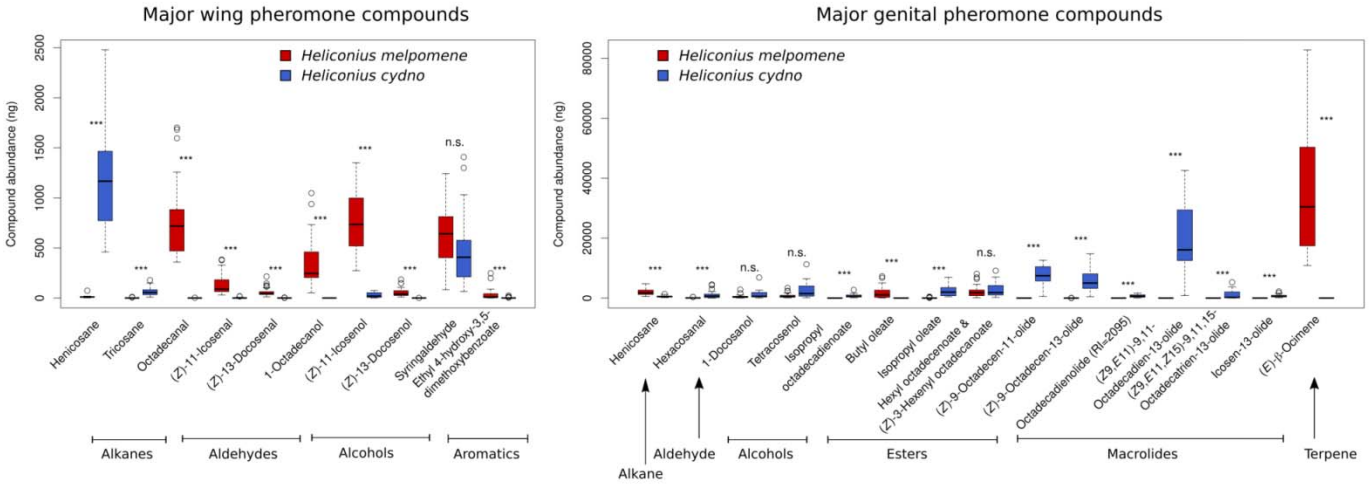
661 Wang, H.-L., Liénard, M. A., Zhao, C.-H., Wang, C.-Z., & Löfstedt, C. (2010) Neofunctionalization in an
662 ancestral insect desaturase lineage led to rare Δ6 pheromone signals in the Chinese tussah silkworm.
663 Insect Biochemistry and Molecular Biology, 40, 742-751. http://doi: 10.1016/j.ibmb.2010.07.009

664 Wiley, C., Ellison, C. K., & Shaw, K. L. (2011) Widespread genetic linkage of mating signals and
665 preferences in the Hawaiian cricket *Laupala*. Proceedings of the Royal Society of London B: Biological
666 Sciences, 279, 1203–1209. http://doi: 10.1098/rspb.2011.1740

667 Wu, W., Cottrell, C. B., Hansson, B. S., & Löfstedt, C. (1999) Comparative study of pheromone production
668 and response in Swedish and Zimbabwean populations of turnip moth, *Agrotis segetum*. Journal of
669 Chemical Ecology, 25, 177-196. http://doi: 10.1023/A:1020849419193

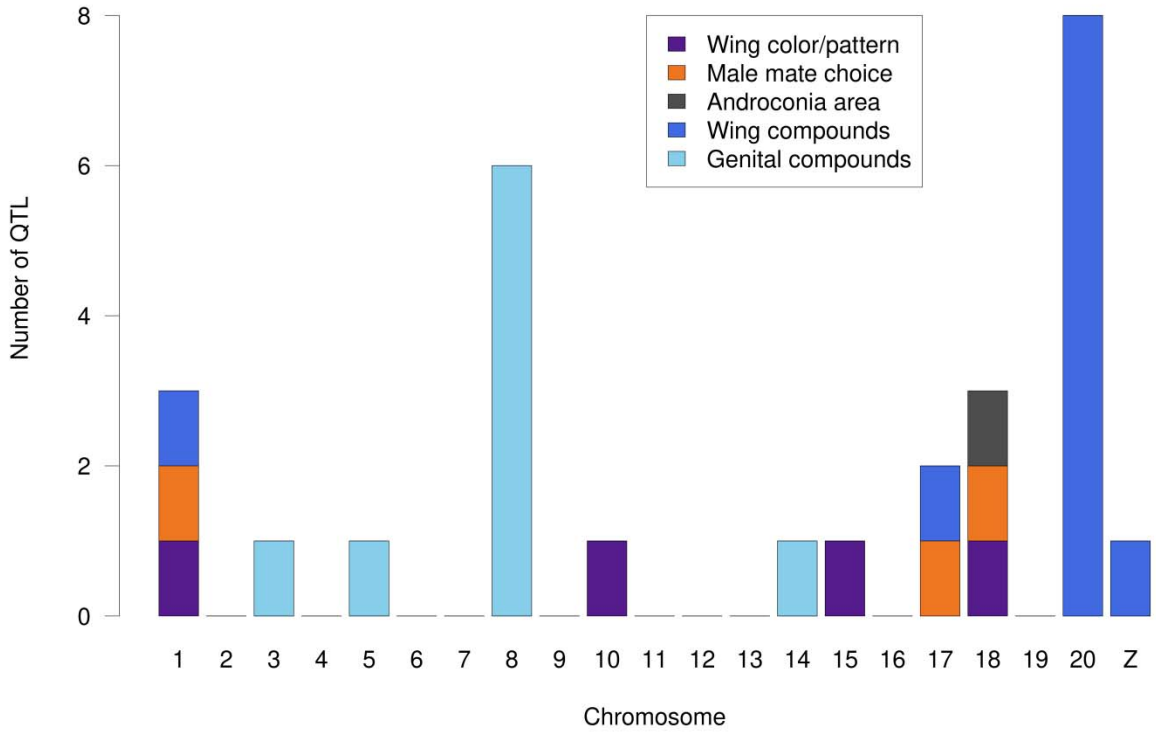
670 **Figures**

671 Figure 1: Major compounds found in the wings (left) and genitals (right) of *Heliconius melpomene* and *H.*
672 *cydno*, arranged by compound type. Only those compounds comprising at least 1% of either species'
673 pheromone bouquet are included. n.s., not significant; ***: $p < 0.005$.



674
675

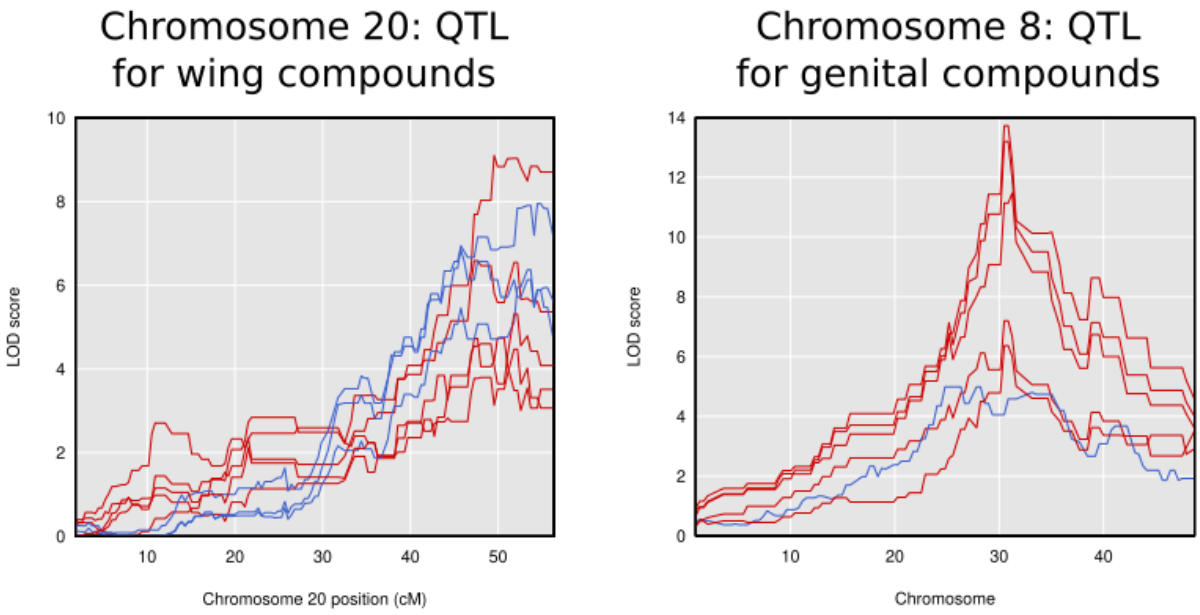
676 Figure 2: QTL for various traits across the *Heliconius melpomene* and *H. cydno* genomes. Wing
677 color/pattern loci represent the four major color loci (Jiggins, 2017); male mate choice are from (Merrill
678 et al., 2019); (relative) androconia area and Bonferroni-significant wing and genital compounds from this
679 study.



680

681

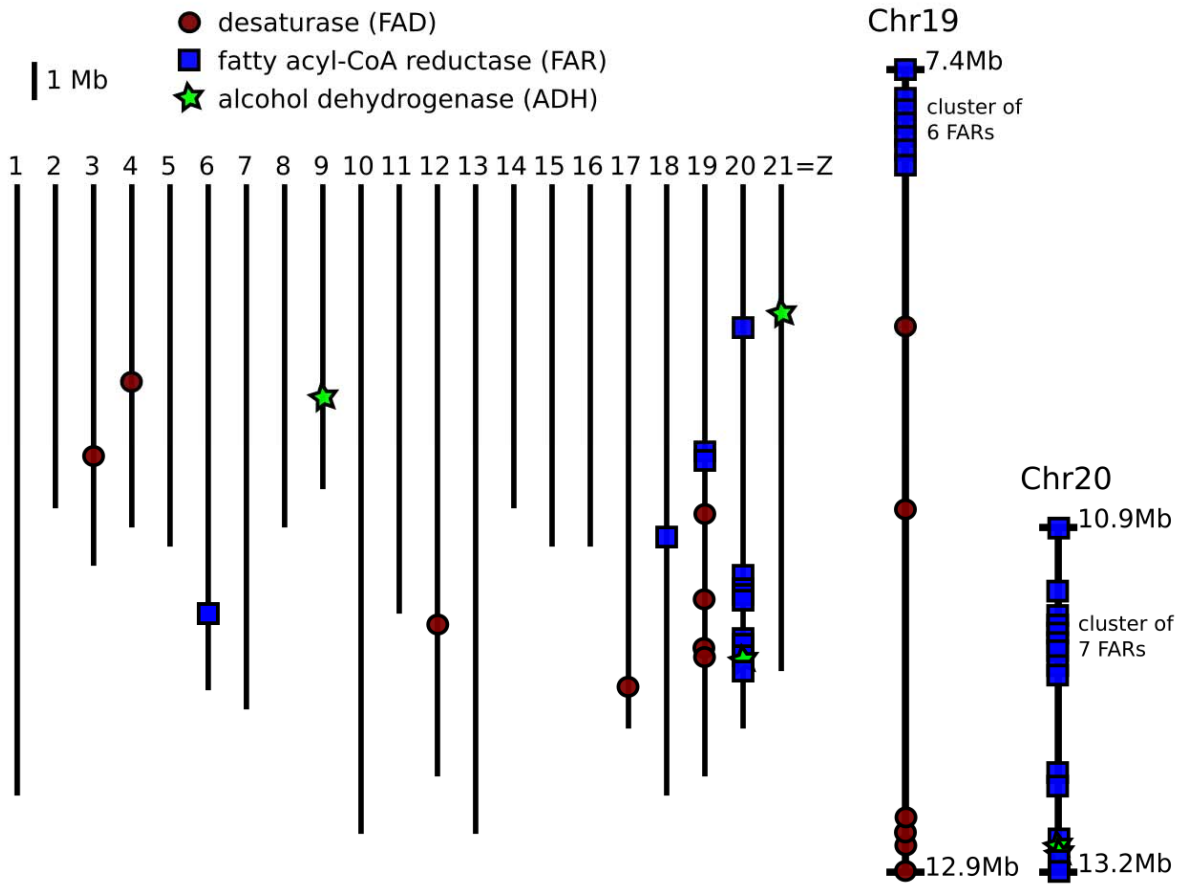
682 Figure 3: Chromosome-specific QTL plots for compounds with significant QTL (after Bonferroni
683 correction) on those chromosomes. Red: compound mapped in backcrosses to *H. melpomene*; blue:
684 compound mapped in backcrosses to *H. cydno*.



685

686

687 Figure 4: Location of candidate genes for fatty acid-derived compound biosynthesis across the genome
688 of *Heliconius melpomene*.



689

690

691 **Tables**

692 Table 1: Details of compounds with Bonferroni-significant QTL in *Heliconius* wing androconia. a: from
693 (Byers et al., 2020). c: in backcross to *H. cydno*. m: in backcross to *H. melpomene*.

Compound	**Type**	**Mean (*H. melpomene*, ng)[a]**	**Mean (*H. cydno*, ng)[a]**	**Bonferroni-significant QTL on chromosome**	**QTL position on chromosome (bp)**
(*Z*)-9-Octadecenal	Unsaturated aldehyde	10.94	Absent	1[m]	9955174[m]
Benzyl cyanide	Aromatic (nitrogenous)	Absent	1.813	17[m]	5596551[m]
Octadecanal[a]	Saturated aldehyde	783.4	0.569	20[c]; 20[m]	10947362[c]; 11725848[m]
Henicosane	Alkane	12.45	1152	20[m]	12126129[m]
Methyloctadecanol	Saturated alcohol	12.23	Absent	20[m]	12374765[m]
1-Octadecanol[a]	Saturated alcohol	345.6	Absent	20[m]	12652775[m]
(*Z*)-16-Methyl-9-octadecenol	Unsaturated alcohol	7.501	Absent	20[m]	12652775[m]
(*Z*)-13-Docosenal	Unsaturated aldehyde	62.56	0.276	20[c]	12917911[c]
(*Z*)-11-Icosenal	Unsaturated aldehyde	150.7	3.290	20[c]	13267611[c]
Methyl 3-indoleacetate	Aromatic (nitrogenous)	3.782	0.013	Z[c]	8478732[c]

694

695 Table 2: Details of compounds with Bonferroni-significant QTL in *Heliconius* genitals. a: from (Darragh et
696 al., 2020). c: in backcross to *H. cydno*. m: in backcross to *H. melpomene*.

Compound	**Type**	**Mean (*H. melpomene*, ng)[a]**	**Mean (*H. cydno*, ng)**	**Bonferroni-significant QTL on chromosome**	**QTL position on chromosome (bp)**
Hexyl 3-methylbutyrate	Saturated ester	Absent	421.1	3[m]	234227[m]
2-*sec*-Butyl-3-methoxypyrazine	Aromatic (nitrogenous)	32.66	3.569	5[c]	4427722[c]
12-Octadecanolide	Macrolide	Absent	105.4	8[c]	4512378[c]
(*Z*)-9-Octadecen-11-olide	Macrolide	Absent	7509	8[m]	5341821[m]
(*Z*9,*E*11)- 9,11-Octadecadien -13-olide	Macrolide	Absent	19714	8[m]	5341821[m]
Hexadecenolide	Macrolide	6.866	165.5	8[m]	5341821[m]
Hexyl octadecanoate	Saturated ester	121.8	24.21	8[m]	5987924[m]
(*Z*)-9-Octadecen-13-olide	Macrolide	7.282	5856	8[m]	6072562[m]
Isopropyl oleate	Unsaturated ester	50.22	2259	14[m]	7370836[m]

697

698

Supplemental Figures and Tables

Supplemental Figure 1: QTL plots for all mapped wing compounds. Dashed horizontal line: p = 0.05 significance threshold. Solid horizontal line: Bonferroni-corrected significance threshold. BC, backcross.

Supplemental Figure 2 (in two parts): QTL plots for all mapped genital compounds. Dashed horizontal line: p = 0.05 significance threshold. Solid horizontal line: Bonferroni-corrected significance threshold. BC, backcross.

Supplemental Figure 3: QTL plots for relative hindwing androconial area. Solid horizontal line: p = 0.05 significance threshold. HW, hindwing. BC, backcross.

Supplemental Figure 4: Whole-genome QTL plots for compounds with at least one significant QTL. Red: compound mapped in backcrosses to *H. melpomene*; blue: compound mapped in backcrosses to *H. cydno*. A: Compounds with at least one QTL significant at the p = 0.05 level shown. B: Compounds with at least one QTL significant after Bonferroni correction shown.

Supplemental Table 1: Full list of compounds found in the wings and genitalia of *Heliconius melpomene* and *H. cydno*, details of statistical comparisons of amounts in both species, and cross direction(s) the compound was mapped in, if any.

Supplemental Table 2: Details of quantitative trait loci passing a significance threshold of either p = 0.05 or a Bonferroni-corrected threshold.