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

Kelsey J. R. P. Byers[1,2,a*], Kathy Darragh[1,2,b,*], Sylvia Fernanda Garza[2,c], Diana Abondano Almeida[2,d], 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 Cell and Developmental Biology, John Innes Centre, Norwich, UK*

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

[c] *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*

[d] *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

email K.J.R.P.B.: Kelsey.Byers@jic.ac.uk
email K.D.: kdarragh@ucdavis.edu
email S.F.G.: sylvia.f.garza@gmail.com
email D.A.A.: AbondanoAlmeida@bio.uni-frankfurt.de
email I.A.W.: iaw22@cam.ac.uk
email P.M.A.R.: pasi.rastas@helsinki.fi

email R.M.M.: merrill@biologie.uni-muenchen.de
email S.S.: stefan.schulz@tu-bs.de
email W.O.M.: mcmillano@si.edu
email C.D.J.: c.jiggins@zoo.cam.ac.uk

Running title: Clustering of pheromone loci in *Heliconius*

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

Author contributions:

KJRPB and KD should be considered joint first author. CDJ, KD, and KJRPB 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 the role of endogenously-produced compounds 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 male 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 (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 endogenously-produced or -modified 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. Both desaturases and fatty acyl-CoA reductases appear to be commonly involved in pheromone biosynthetic variation. Desaturases introduce double or triple bonds between carbon molecules in pheromone components, for example turning alkanes (less common as pheromone components) into alkenes (a common component of lepidopteran pheromones). Fatty acyl-CoA reductases, or FARs, turn fatty acids into fatty alcohols, for example turning octadecanoic acid into octadecanol. 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 are responsible for

117 species-specific pheromone alterations in *Ostrinia nubialis* strains (Lassance et al., 2010, 2013). Within
118 single species, desaturases and FARs have received the most attention, being functionally characterized
119 in a variety of Lepidoptera including *Yponomeuta* (Liénard et al 2010), *Agrotis* (Ding & Löfstedt, 2015),
120 *Manduca* (Buček et al., 2015), *Bombyx* (Moto et al., 2004), and *Antheraea* (Wang et al., 2010). In some
121 cases, variation in these female pheromones have also been shown to contribute to reproductive
122 isolation (e.g. Liebherr & Roelofs, 1975, Wu et al., 1999, Emelianov et al., 2001; reviewed in Smadja &
123 Butlin, 2009).

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

134 Speciation is typically thought to rely on the accumulation of multiple reproductive barriers, involving
135 divergence in many different traits (Butlin & Smadja, 2017). In order to understand the extent of genetic
136 linkage that underlies species differences it will therefore be useful to study taxa in which multiple
137 different traits can be mapped. *Heliconius* butterflies have been well studied in the context of speciation
138 (Bates, 1861, Merrill et al., 2015, Jiggins, 2017), and in particular we know a great deal about the species
139 complex of *Heliconius melpomene* and its sister lineage *Heliconius cydno/timareta*. Genetic loci
140 underlying wing pattern and mate preference have been mapped in this group (Merrill et al., 2010,
141 Jiggins, 2017, Merrill et al., 2019). The additional role of chemical signaling in reproductive isolation has
142 long been suspected (Jiggins, 2008), but only recently studied in any detail. Chemical profiles of the male
143 wing androconia (patches of specialist scales on male wings that release pheromones) and genital
144 regions (thought to act as aphrodisiacs and anti-aphrodisiacs, respectively) differ between species
145 (Estrada et al., 2011, Mann et al., 2017, Darragh et al., 2020), and are important for mate choice,
146 including altering behavior towards con- and heterospecific individuals (Mérot et al., 2015, Darragh et
147 al., 2017, González-Rojas et al., 2020). Genital pheromones play a more complex role in mate choice in

148 *Heliconius*, as they are transferred by males to females during mating and subsequently decrease
149 advances by other males (Gilbert, 1976, Schulz et al., 2007, Estrada, 2009).

150 One challenge of studying the entire chemical profile is determining which (combination of) compounds
151 are behaviorally active. The chemical bouquets of wings and genitals are complex, often consisting of
152 30-70 compounds (Mann et al., 2017, Darragh et al., 2017, 2020, Byers et al., 2020, this study), but it
153 remains unknown how many of these are important for signaling. A single wing compound in *H.*
154 *melpomene*, octadecanal, is known to be biologically active in *H. melpomene* and *H. cydno* (Byers et al.,
155 2020), while the genital compound (*E*)-β-ocimene acts as the main anti-aphrodisiac in *H. melpomene*
156 (Schulz et al., 2007). In addition, one bioactive genital compound in *H. cydno* (hexyl 3-methylbutyrate,
157 Estrada, 2009) is known. It seems likely that additional compounds in the complex bouquet (30-70
158 compounds) produced by each of these species are biologically active, but this remains to be tested.
159 The simple genetic control of pheromone components in other systems and their role as secondary
160 metabolism products suggests that they can be altered relatively simply without major pleiotropic
161 consequences to essential organismal functions, as has been seen in some moth species (Groot et al.,
162 2019), though pleiotropic effects have been seen in *Drosophila* (Bousquet et al., 2012 and Zelle et al.,
163 2020). In *Drosophila* some loci specifically showed pleiotropy of production and perception of
164 pheromones, thought to be rare in Lepidoptera (Haynes, 2016). In combination with their importance in
165 inter- and intraspecific mate choice, such simple genetic control and relatively lower risk of pleiotropy
166 makes them ripe material for adaptation and speciation. We here conduct quantitative trait locus (QTL)
167 analyses for wing androconial area and wing and genital compounds that differ between *Heliconius*
168 *melpomene* L. and *H. cydno* Doubleday (Figure 1; see also Byers et al., 2020, Darragh et al., 2020). We
169 then investigated the patterns of QTL distribution across the genome to test for clustering of loci across
170 chromosomes.

171

172 **Methods**

173 *Heliconius melpomene rosina* and *H. cydno chioneus* (hereafter *H. melpomene* and *H. cydno*) and their
174 interspecific hybrids were reared in outdoor insectaries at the Smithsonian Tropical Research Institute in
175 Gamboa, Panama. Initial stocks were established from wild-caught individuals, and outbred stocks were
176 refreshed with more wild-caught individuals as necessary. Sexes were kept separately and virgin females
177 were kept separate from egg-laying females. To control for host plant effects, all larvae were fed on the

178 same host plant (*Passiflora platyloba* var. *williamsi*), and adults were maintained on a 20% sugar
179 solution and *Gurania eriantha*, *Psiguria triphylla*, *Psiguria warscewiczii*, and *Psychotria poeppigiana* as
180 pollen sources. Pheromones were collected from male individuals (both parental species and offspring
181 of mapping crosses) following (Darragh et al. 2017). Wing and genital data from *H. melpomene* and wing
182 data from *H. cydno* have previously been published in (Darragh et al., 2017, Darragh et al., 2019, Byers
183 et al., 2020). Briefly, males of 7-14 days post eclosion had their wings and genitals removed and soaked
184 in dichloromethane plus 1 ng/µL 2-tetradecyl acetate (DCM+IS) for one hour, after which the DCM+IS
185 was removed to a fresh vial and stored at -20°C before gas chromatography-mass spectrometry analysis.
186 Bodies were collected and preserved in 20% dimethyl sulfoxide (DMSO) and 0.25 M EDTA (pH 8.0) and
187 stored at −20°C for later DNA extraction.

188 Wing and genital extractions were run on an Agilent 7890B gas chromatograph coupled with an Agilent
189 5977 mass spectrometer (GC-MS) with electron ionization (Agilent Technologies, California, USA) with
190 an HP-5MS capillary column (30m length, 0.25mm inner diameter) and helium as the carrier gas (1.2
191 mL/min). Injection used an Agilent ALS 7693 autosampler and was splitless, with an inlet temperature of
192 250°C. The oven was first held isothermally at 50°C for five minutes, then increased at 5°C/minute to
193 320°C and again held isothermally for five minutes. Compounds were identified with a custom MS
194 library and quantified using the internal standard area. Only compounds that could be definitively
195 identified were included in the data set. To be included in the mapping experiment, compounds had to
196 be present in at least two thirds of the individuals from one of the parent species, i.e. they had to be
197 "typical" *H. melpomene* or *H. cydno* compounds. A limited set of wing compounds (three
198 methyloctadecanals and two henicosenes from *H. melpomene* and *H. cydno* respectively) were excluded
199 from the wing analysis due to close retention times in the GC-MS data preventing definitive
200 identification of the species phenotype in the mapping individuals. Absolute androconial area and
201 absolute hindwing area were measured in the GNU Image Manipulation Program (GIMP) as pixel counts,
202 then absolute androconial area was divided by hindwing area to produce the percentage of the
203 hindwing taken by the androconial region, the relative androconial area.

204 Mapping crosses consisted of backcrosses of either an *H. cydno* or *H. melpomene* mother and an F_1
205 father, as F_1 females are normally sterile. Ten families were constructed as backcrosses to *H. melpomene*
206 (representing 89 individuals for the wing phenotype, 81 individuals for the genital phenotype, and 78
207 individuals for the hindwing relative androconial area phenotype) and fifteen families were constructed
208 as backcrosses to *H. cydno* (127 individuals for the wing, 114 individuals for the genital phenotype, and

209 124 individuals for the hindwing relative androconial area phenotype) for a total of 216 and 195 hybrid
210 individuals for wing and genital studies respectively. DNA extraction, library preparation, and QTL map
211 construction are detailed in (Byers et al., 2020, Darragh et al., bioRxiv). Briefly, Qiagen DNeasy kits
212 (Qiagen) were used for DNA extraction, and individuals then genotyped either by RAD-seq or low-
213 coverage whole genome sequencing using nextera-based libraries (Picelli et al., 2014, Davey et al., 2017,
214 Merrill et al., 2019). Samples were sequenced by HiSeq 3000 (Illumina) by BGI (China). Linkage mapping
215 was conducted using the standard Lep-MAP3 (LM3) pipeline (Rastas, 2017). The initial linkage groups
216 and marker order were constructed based on the *H. melpomene* genome for 21 chromosomes, as *H.*
217 *melpomene* and *H. cydno* have highly colinear genomes (Davey et al., 2017). This resulted in a linkage
218 map with 447,818 SNP markers, which was then evenly thinned by a factor of ten to 44,782 markers to
219 ease computation.

220 To determine which cross direction to use for each compound, we inspected the distribution of the
221 compound phenotype in the parents, F_1 individuals, and the backcross individuals. Compounds that
222 were present in equal amounts in both parent species (assessed using a Kruskal-Wallis test) were not
223 mapped. Once cross direction was determined, the remaining compounds were log-transformed to
224 approximate normality before being regressed against the linkage map using R/qtl2 (Broman et al.,
225 2018). Due to the family structure present in our crosses, we additionally included a kinship matrix
226 calculated by R/qtl2 using the LOCO (leave one chromosome out) method. Permutation testing was
227 used (with 1000 replicates) to determine QTL threshold. Compounds are likely to be correlated and our
228 data set represents multiple testing of the same mapping crosses, therefore, we used Bonferroni
229 correction separately for each cross direction to establish a second, more conservative, LOD threshold
230 for significance by dividing the threshold by the number of compounds (traits) mapped in the relevant
231 backcross direction. No correction was used for relative androconial area mapping as the trait was
232 mapped alone due to its general lack of correlation with individual volatile compounds. QTL confidence
233 intervals were obtained using bayes_int. As R/qtl2 does not provide a function to calculate the
234 traditional percentage of variance explained by a specific marker, we instead calculated the percentage
235 of the parental difference explained by the genotype at a given marker as a fraction with the difference
236 between the average phenotype of the two genotypes as the numerator and the difference between
237 the average phenotype of the two parental species as the denominator. This can result in values over
238 100% when more variance is present in the mapping population than the difference between the parent
239 species accounts for.

240 To identify clustering of QTL, we first calculated the expected number of QTL per chromosome, taking
241 into account the chromosome's length in cM, then compared that against the observed distribution of
242 QTL per chromosome using a chi-squared test, computing p-values with Monte Carlo permutation with
243 B = 10,000 permutations. Examination of the residuals was used to determine which chromosome(s)
244 displayed clustering, with any residual over 2 considered significant, as in (Erickson et al., 2016). We also
245 repeated these tests taking into account chromosome length in basepairs because the relationship
246 between map length and physical length can differ across chromosomes. The results were unchanged.

247 As our linkage map is based on whole genome sequencing data, we were able to recover the basepair
248 intervals corresponding to the QTL centimorgan (cM) confidence intervals (defined as the outermost
249 markers at those cM positions) as well as their peaks (defined as the first marker in at that cM position).
250 Lepbase (Challis et al., 2016) was queried to identify genes within the main interval on chromosome 20
251 for wing fatty acid-derived compound production (identified as the minimum overlapping window of all
252 Bonferroni-significant compound confidence intervals). Genes were searched against the nr (non-
253 redundant) protein database using BLASTp (Altschul et al., 1990) to obtain putative functional
254 annotations, with no specific cutoff used to define a putative annotation; annotations were made when
255 multiple hits with high alignment scores agreed on putative function.

256

257 **Results**

258 We first looked at the distribution of compounds in *H. melpomene*, *H. cydno*, and their hybrids. Analysis
259 of the pheromones found a total of 31 compounds in the wings and 68 compounds in the genitals across
260 the two species that were possible to map (i.e. not unknown or overlapping in the gas chromatography-
261 mass spectrometry trace) (Figure 2, Table 1). There was limited overlap in compounds between the
262 wings and genitals, with only the straight-chain alkanes henicosane, docosane, tricosane, pentacosane,
263 and hexacosane (all of which may be part of the normal whole-body cuticular hydrocarbon profile), the
264 unsaturated aldehyde (*Z*)-11-icosenal, and the nitrogenous aromatic benzyl cyanide found in both body
265 regions. A subset of compounds (6 in the wings and 18 in the genitals) were not mapped as they did not
266 differ significantly between the parental species. The rest showed significant differences and thus had
267 adequate variation for QTL mapping to be feasible. Of the wing compounds, 11 showed segregation in
268 backcrosses to *H. cydno* and thus were mapped with those individuals; nine were mapped in
269 backcrosses to *H. melpomene*; and five showed segregation in both backcross directions and thus were

270 mapped separately in backcrosses to each species. For the genital compounds, 19 were mapped in
271 backcrosses to *H. cydno*; 17 in backcrosses to *H. melpomene*; and 14 in both directions.

272 Most compounds mapped did not produce significant QTL at either the p = 0.05 or the Bonferroni-
273 adjusted cutoff levels (the latter taking into account the number of compounds mapped) (Figures S1 and
274 S2). A total of 21 genital compounds and 14 wing compounds produced significant QTL, including three
275 peaks already published in (Byers et al., 2020, Darragh et al., bioRxiv) for octadecanal, 1-octadecanol,
276 and (*E*)-β-ocimene. Of these QTL, nine genital and ten wing compounds were significant at the more
277 conservative Bonferroni threshold (Tables 2 and 3; Table 4). The higher drop-out rate for genital
278 compounds after Bonferroni correction is likely the result of the more stringent correction due to a
279 higher number of compounds corrected for. We also found a single significant QTL on chromosome 18
280 for relative hindwing androconial area in backcrosses to *H. cydno* (Figure S3). The percentage of parental
281 difference explained by the peak markers at these QTL was highly variable, with values ranging from 0.6-
282 130% for wing compounds and 1.9-170% for genital compounds, but most compounds fell within a
283 range of approximately 20-50% of parental difference explained by the locus. Surprisingly, there was no
284 obvious association between the amount of variance present in the backcrosses and the presence or
285 absence of a QTL for each compound, as seen from the coefficient of variance values in Table 1, which
286 did not differ between compounds that did and did not have a QTL associated (t = 0.342, df = 78, p =
287 0.73).

288 Compounds that showed significant QTL tended to be of specific chemical classes. For the wings, of the
289 14 wing compounds with significance at p = 0.05, 12 were fatty acid-derived compounds (FADs),
290 including alkanes as well as oxygenated alkanes and alkenes (aldehydes and alcohols). The remaining
291 two significant wing compounds were both nitrogenous aromatics. For the genitals, 10 compounds
292 producing significant QTL were macrolides and eight were FADs (two alkanes, four saturated esters, and
293 two unsaturated esters). The remaining three compounds were again nitrogenous aromatics (two
294 compounds) and one terpene. This is in general agreement with the distribution of compound types in
295 both body regions. The mapped wing compounds comprise 18 fatty acid-derived compounds, three
296 nitrogenous aromatics, two non-nitrogenous aromatics, and two terpenoids, so the finding of significant
297 QTL mostly for fatty acid-derived compounds is not surprising. For the genitals, the mapped compounds
298 consist of 22 fatty acid-derived compounds, 21 macrolides, four terpenoids, two nitrogenous aromatics,
299 and one lactone.

300 Significant quantitative trait loci showed some same-chromosome clustering on specific chromosomes
301 (Figure S4). Notably, these clusters did not overlap with chromosomes harboring known QTL for color
302 pattern, mate choice, or other traits, which generally tend not to be clustered in the genome, apart from
303 a close association between mate choice and a wing patterning gene (Figure 3). When either all QTL
304 significant at $p < 0.05$ or Bonferroni-corrected significant QTL were combined with known QTL,
305 clustering at the chromosomal level was statistically significant ($\chi^2 = 64.036$ for Bonferroni-significant
306 QTL, $\chi^2 = 78.524$ for all significant QTL, $p < 0.001$ for both) for chromosomes 8 and 20. In particular, all
307 but one QTL for FAD compounds in the wings mapped to chromosome 20 (the exception mapped to
308 chromosome 1 as well as having a peak on chromosome 20). The other significant wing compounds (all
309 nitrogenous aromatics) mapped to chromosomes 10, 17, 19, and the Z chromosome (chromosome 21).
310 QTL on chromosome 20 were broadly overlapping, in part due to our relatively small mapping
311 populations producing broad Bayesian confidence intervals for QTL location (Figure 4). In addition, the
312 peaks themselves were strongly overlapping on the latter half of chromosome 20. This chromosomal-
313 level clustering was statistically significant ($\chi^2 = 178.39$ for all significant QTL, $\chi^2 = 113.02$ for Bonferroni-
314 corrected significant QTL, $p < 0.001$ in both cases) and residual analysis identified chromosome 20 as the
315 sole outlier. When broken down by chemical class, only FAD compounds clustered significantly at the
316 chromosomal level, again on chromosome 20 ($\chi^2 = 232.49$ for all significant QTL, $p < 0.001$), and this was
317 also significant when only Bonferroni-corrected significant QTL were tested ($\chi^2 = 137.55$, $p < 0.001$).

318 Similar compound class-specific clustering at the chromosome level was seen to a lesser extent for the
319 genitals, where six of the ten macrolides had QTL on chromosome 8, with the others more broadly
320 dispersed on chromosomes 2, 7, 13, 14, and 20. The fatty acid-derived compounds (mostly esters)
321 mapped more broadly across the genome than in the wings, with QTL on chromosomes 1, 2, 3, 6, 8, 9,
322 14, and 20. Chromosome-level clustering was again statistically significant ($\chi^2 = 46.037$ for all significant
323 QTL, $\chi^2 = 101.83$ for Bonferroni-corrected significant QTL, $p = 0.0017$ and $p < 0.001$ respectively) and
324 residual analysis identified chromosomes 8 and 14 as the outliers when all QTL were considered, and
325 chromosome 8 when only Bonferroni-corrected QTL were included. The clustering on chromosome 8 is
326 solely due to the presence of macrolide QTL on this chromosome, while chromosome 14 contains QTL
327 for both a macrolide and three FAD compounds. When broken down by chemical class, only macrolide
328 compounds clustered significantly at the chromosomal level in the genitals, again on chromosome 8 (χ^2
329 $= 58.079$ for all significant QTL, $p < 0.001$), and this was also significant when only Bonferroni-corrected
330 significant QTL were tested ($\chi^2 = 110.88$, $p < 0.001$). Again, the QTL peaks themselves were strongly
331 overlapping in the middle of chromosome 8 (Figure 4) rather than being dispersed across the

332 chromosome. Two compounds had significant QTL in both wings and genitals: henicosane and benzyl
333 cyanide. Both data sets showed QTL on the same chromosomes (20 and 17 respectively, with additional
334 QTL for benzyl cyanide in the wings on other chromosomes). The henicosane QTL were within 1.5cM of
335 one another, while the benzyl cyanide QTL on chromosome 17 were 2.3cM apart.

336 The strong clustering of QTL for wing fatty acid-derived compounds on chromosome 20 suggested a
337 common gene or set of linked genes responsible for fatty acid-derived compound metabolism or its
338 regulation. We searched the region between 43.11 - 56.37 cM (9,398,348 - 14,585,564 basepairs) for
339 potential candidate genes that could be responsible for this clustering. This region spanned the
340 confidence intervals for all Bonferroni-significant QTL apart from 1-octadecanol and contained 363
341 genes. Annotated proteins from this region (from Lepbase, Challis et al., 2016) were searched using
342 BLASTp against the NCBI non-redundant (nr) database, revealing a total of 20 genes potentially involved
343 in the biosynthesis of the wing compounds within this confidence interval on chromosome 20. These
344 genes include two acetyl-CoA carboxylases (involved in primary metabolism), fifteen fatty acyl-CoA
345 reductases (FARs), and two alcohol dehydrogenases (the latter two families involved in both primary
346 and secondary metabolism), none of which have been functionally characterized. Comparing our
347 candidates to those identified in (Liénard et al., 2014), none of our FARs fall within the pheromone gland
348 FAR (pgFAR) clade, whose members fall on chromosome 19. Inspecting the genome for additional
349 secondary metabolism fatty acid-derived metabolism genes, we identified clusters of FARs on
350 chromosomes 19 and 20 (Figure 5), with a more even distribution of alcohol dehydrogenases and
351 desaturases across the genome. Notably, we did not identify any QTL on chromosome 19 for fatty acid-
352 derived compound production in the wings.

353

354 **Discussion**

355 The genetic basis for traits contributing to reproductive isolation remains poorly studied. This is
356 especially the case for differences in chemical signaling between closely related non-model species, in
357 particular for species with complex chemical signatures where multiple components may play a role in
358 mate choice or species recognition. Here, we mapped a total of 40 QTL controlling the segregation of
359 wing and genital pheromone bouquet compounds in crosses between *Heliconius melpomene* and its
360 close relative *H. cydno*. There was significant evidence for non-random clustering of QTL, most notably
361 for wing compounds on chromosome 20, which showed broadly overlapping QTL with closely spaced

362 peaks. There could be a number of reasons for this clustering of loci. First, theory predicts that
363 mutations controlling species differences may accumulate, and persist, in regions already resistant to
364 gene flow, such that recombination is less likely to disrupt species differences (Smadja & Butlin, 2011).
365 Second, functionally similar genes are likely to occur in close linkage due to *e.g.* tandem gene
366 duplication, such that chemically similar compounds could be regulated by linked genes due to
367 functional similarities between their synthesis pathways (Osbourn, 2010). Finally, related compounds
368 could differ between species as a pleiotropic effect of the same mutation if they are downstream
369 products of the same biochemical pathway or in the case of a multifunctional enzyme. The end result of
370 all of these processes is linkage that can facilitate speciation by impeding recombination that will
371 otherwise disrupt the genetic associations between traits that characterize emerging species.

372 In the case of the compounds described here, it is likely that shared biosynthetic origins play an
373 important role in their chromosomal distributions. The majority of the wing compounds found in either
374 species are fatty acid-derived compounds (FADs; 19 of 31), and the majority of these FADs (12 of 19)
375 were associated with a significant QTL on chromosome 20. It is interesting that there was clustering of
376 one class of biosynthetic enzyme in this pathway, the fatty acyl-CoA reductases (FARs), on both
377 chromosomes 19 and 20, but no QTL on chromosome 19 for any wing or genital FAD compound,
378 suggesting the FARs on chromosome 19 are not involved in putative pheromone compound production.
379 Similarly, QTL for genital compounds exhibited slight clustering of loci underlying macrolide compound
380 production on chromosome 8, but otherwise were more evenly spread across the genome, likely due to
381 their greater biosynthetic diversity. The QTL for shared macrolide compounds may also be due to
382 shared biosynthetic origins of these compounds, but the biosynthesis of macrolides is less well
383 understood and consequently we were unable to identify candidate genes on this chromosome. We also
384 found a QTL on chromosome 3 for hexyl 3-methylbutyrate, a known bioactive component of the genital
385 pheromone of *H. cydno* (Estrada, 2009) whose genetic basis was previously unknown. Of note, the QTL
386 for relative androconial area roughly peaks on the gene *optix* (chromosome 18), a locus controlling red
387 color patterns in *Heliconius* (Reed et al., 2011, Van Belleghem et al., 2017). This locus is also known to
388 control male-specific scales in the basal heliconiine *Dryas iuila* and has an expression domain in the
389 hindwing androconia region (Martin et al., 2014). This overlap raises the possibility for coupled control
390 of color pattern and androconial pheromone production differences between *Heliconius* species.

391 Overall, we found no evidence for linkage of any of the chemical QTL clusters with other loci or QTL that
392 have been shown to play a difference in the strong assortative mating observed between *H. melpomene*

393 and *H. cydno*. The major wing patterning genes are located on chromosomes 1 (*aristalless*), 10 (*WntA*),
394 15 (*Cortex*), and 18 (*optix*), only chromosome 1 of which harbors a chemical QTL, but the peak does not
395 include the *aristalless* gene. In addition, major QTL for male mate choice map to chromosome 1, 17, and
396 18 (Merrill et al., 2019), but, similarly, the chemical QTL we identified on chromosomes 1 and 17 do not
397 overlap with the major QTL identified for male mate choice. We might expect an overlap between
398 chemical QTL and female mate choice QTL (as females likely assess males partially based on their wing
399 androconial chemistry), but the latter has not been mapped in this species pair. Nonetheless, the
400 clustering of wing compounds in particular might contribute to pheromone-based speciation,
401 particularly in the face of gene flow (Via, 2012). If more than one pheromone component that maps to
402 chromosome 20 is involved in reproductive isolation, and if genes responsible for their production are
403 tightly linked (or if pleiotropy occurs), this may also facilitate speciation (Felsenstein, 1981, Smadja &
404 Butlin, 2011, Merrill et al., 2019) and prevent the breakup of potentially adaptive pheromone bouquet
405 mixtures. Evidence suggests that mixture effects are important in pheromone processing (Riffell et al.
406 ,2009, Clifford & Riffell, 2013, Lei et al., 2013), and these mixtures might be maintained through tight
407 genetic linkage between biosynthetic loci, facilitated by gene evolutionary events such as tandem
408 duplication. In contrast to the tight clustering seen on chromosome 20 for wing FAD compounds, we
409 saw more limited clustering for the more chemically diverse genital compounds, likely due to their more
410 diverse potential biosynthesis pathways. This might hamper any potential role in reproductive isolation.
411 Given their known function as anti-aphrodisiacs, genital compounds are not so likely to play a major role
412 in isolation between *Heliconius* species (Gilbert, 1976, Schulz et al., 2007, Estrada, 2009).

413 Some of the compounds found on wings and in genitalia might have a plant origin. Several compounds
414 in the wings and genitals did not differ between parental species, and many of these have been shown
415 to be affected by larval diet in *Heliconius melpomene*. Most of the invariant wing compounds are
416 aromatics, which are expected to be plant derived in *Heliconius* (Darragh et al., 2019). Notably, we
417 reared all mapping butterflies on the same larval and adult diets, thus effectively creating a dietary
418 common garden experiment, and so we do not expect larval or adult diet to affect the result of the QTL
419 mapping. Many male moth pheromone components are derived from host plant volatiles with minimal
420 alterations, potentially due to females' pre-existing biases for detecting these compounds during the
421 host plant search (Conner & Iyengar, 2016). However, many of our mapped *Heliconius* compounds do
422 appear to have a genetic basis (i.e. a QTL was found), in line with the biosynthesis of FAD and related
423 compounds by female moths. Breaking our compounds down into four classes (FADs, macrolides,
424 aromatics, and terpenoids), we failed to find QTL (at the $p = 0.05$ level) for 54%, 58%, 43%, and 83% of

425 these classes respectively across both wings and genitals. In particular, terpenoids that differ between
426 species may be the result of differential larval host plant use. It is of course possible that plant derived
427 compounds could show a genetic basis if there were species differences in the genes regulating their
428 uptake, and one example might be icosanal, which has been shown to be influenced by larval host plant
429 (Darragh et al., 2019) but which is also associated with a QTL in this study.

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

445 Single FAR and alcohol dehydrogenase genes are missing from *H. cydno* (Byers et al., 2020), suggesting
446 an obvious mechanism for the striking difference in FAD compounds between the two species; these
447 differences may also be due to e.g. differential regulation of FARs or other biosynthetic genes, or
448 mutations within these genes. Alternately, a single FAR gene or small set of genes may be responsible
449 for the biosynthesis of the majority of FAD compounds in *Heliconius*, similar to the situation found with
450 *eloF* in *Drosophila*, where knockdown of the gene produces widely pleiotropic effects on cuticular
451 hydrocarbon profiles (Combs et al., 2018). Similar effects are seen in *Spodoptera litura*, where RNAi of
452 *FAR17* alone alters levels of the four major sex pheromone components (Lin et al., 2017). Another
453 option, though one we consider unlikely, is that this tight cluster of FARs on chromosome 20 serves as a
454 supergene, allowing multiple phenotypic traits (here individual chemical compounds) to co-segregate
455 and maintaining strong integration of the chemical phenotype. Supergenes have been described for

456 other phenotypic morphs in *Heliconius* (Joron et al., 2011), but we are unaware of mechanisms on
457 chromosome 20 that might contribute to supergene formation and maintenance, for example a
458 chromosomal inversion, which would have likely been evident in our linkage map.

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

474 Understanding the genetic architecture of complex traits can allow us to understand their potential role
475 in speciation. Determining whether the observed clustering QTL on chromosomes 8 and 20 reflect tight
476 linkage of multiple biosynthetic and/or regulatory genes or a single "master" locus responsible for the
477 biosynthesis or regulation of multiple pheromone components will contribute to our understanding of
478 both pheromone biosynthesis and pleiotropy and linkage in general. With genes in hand, CRISPR (a
479 technique that works in *Heliconius*, Livraghi et al., 2018) could be used to alter gene function, allowing
480 confirmation of the role of pheromone components via behavioral assays. As major effect loci (as these
481 QTL appear to be) and tight linkage (as potentially seen here) can both facilitate the speciation process,
482 especially in the face of gene flow (Felsenstein, 1981, Merrill et al., 2010, Smadja & Butlin, 2011, Via,
483 2012), the QTL identified here could be barrier or speciation genes in *Heliconius*. These specific QTL may
484 be involved in interspecific mate choice and reproductive isolation, and thus the genes underlying them
485 could contribute to the maintenance of species boundaries. We have here identified a total of 40 QTL
486 underlying species differences in 33 potential pheromone components, demonstrated clustering of

these QTL for wing and genital compounds on specific chromosomes not linked to known loci for species differences, and identified candidate genes underlying the production of the major chemical class of wing compounds. Together, these findings further our knowledge of chemical ecology, pheromone genetics, and their potential roles in *Heliconius* butterfly mate choice and speciation.

References

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

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

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

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

Bousquet, F., Nojima T., Houot, B., Chauvel, I., Chaudy, S., Dupas, S., …Ferveur, J.-F. (2012) Expression of a desaturase gene, *desat1*, in neural and nonneural tissues separately affects perception and emission of sex pheromones in *Drosophila*. Proceedings of the National Academy of Sciences USA, 109, 249-254. http://doi: 10.1073/pnas.1109166108

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

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

513 Buček, A., Matoušková, P., Vogel, H., Šebesta, P., Jahn, U., Weißflog., J., ...Pichová, I. (2015) Evolution of
514 moth sex pheromone composition by a single amino acid substitution in a fatty acid desaturase.
515 Proceedings of the National Academy of Sciences USA, 112, 12586-12591. http://doi:
516 10.1073/pnas.1514566112

517 Butlin, R. K., & Smadja, C. M. (2017) Coupling, reinforcement, and speciation. The American Naturalist,
518 191, 155-172. http://doi: 10.1086/695136

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

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

524 Clifford, M. R., & Riffell, J. A. (2013) Mixture and odorant processing in the olfactory systems of insects: a
525 comparative perspective. Journal of Comparative Physiology A, 199, 911-928. http://doi:
526 10.1007/s00359-013-0818-6

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

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

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

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

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

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

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

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

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

552 Erickson, P. A., Glazer, A. M., Killingbeck, E. E., Agoglia, R. M., Baek, J., Carsanaro, S. M., ...Miller, C. T.
553 (2016) Partially repeatable genetic basis of benthic adaptation in threespine sticklebacks. Evolution, 70,
554 887-902. http://doi: 10.1111/evo.12897

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

557 Estrada, C., Schulz, S., Yildizhan, S., & Gilbert, L. E. (2011) Sexual selection drives the evolution of
558 antiaphrodisiac pheromones in butterflies. Evolution, 65, 2843-2854. http://doi: 10.1111/j.1558-
559 5646.2011.01352.x

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

562 Fujii, T., Yasukochi, Y., Rong, Y., Matsuo, T., & Ishikawa, Y. (2015) Multiple Δ11-desaturase genes
563 selectively used for sex pheromone biosynthesis are conserved in *Ostrinia* moth genomes. Insect
564 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

Groot, A. T., van Wijk, M., Villacis-Perez, E., Kuperus, P., Schöfl, G., van Veldhuizen, D., & Heckel, D. G. (2019) Within-population variability in a moth sex pheromone blend, part 2: selection towards fixation. Royal Society Open Science, 6, 182050. http://doi: 10.1098/rsos.182050

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

Haynes, K. F. (2016) Genetic control of moth sex pheromone signal and response. In Pheromone Communication in Moths, J. D. Allison and R. Cardé, ed. Oakland, CA, USA: University of California Press.

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

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

594 Joron, M., Frezal, L., Jones, R. T., Chamberlain, N. L., Lee, S. F., Haag, C. R., ...ffrench-Constant, R. H.
595 (2011) Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry.
596 Nature, 477, 203-206. http://doi: 10.1038/nature10341

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

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

601 Lassance, J.-M., Groot, A. T., Liénard, M. A., Antony, B., Borgwardt, C., Andersson, F., ...Löfstedt, C.
602 (2010) Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones.
603 Nature, 466, 486-489. http://doi: 10.1038/nature09058

604 Lassance, J.-M., Liénard, M. A., Antony, B., Qian, S., Fujii, T., Tabata, J., ...Löfstedt, C. (2013) Functional
605 consequences of sequence variation in the pheromone biosynthetic gene *pgFAR* for *Ostrinia* moths.
606 Proceedings of the National Academy of Sciences USA, 110, 3967-3972. http://doi:
607 10.1073/pnas.1208706110

608 Lei, H., Chiu, H.-Y., & Hildebrand, J. G. (2013) Responses of protocerebral neurons in *Manduca sexta* to
609 sex-pheromone mixtures. Journal of Comparative Physiology A, 199, 997-1014. http://doi:
610 10.1007/s00359-013-0844-4

611 Li, R.-T., Ning, C., Huang, L.-Q., Dong, J.-F., Li, X., & Wang, C.-Z. (2017) Expressional divergences of two
612 desaturase genes determine the opposite ratios of two sex pheromone components in *Helicoverpa*
613 *armigera* and *Helicoverpa assulta*. Insect Biochemistry and Molecular Biology, 90, 90-100. http://doi:
614 10.1016/j.ibmb.2017.09.016

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

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

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

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

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

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

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

636 Mann, F., Vanjari, S., Rosser, N., Mann, S., Dasmahapatra, K. K., Corbin, C., ...Schulz, S. (2017) The scent
637 chemistry of *Heliconius* wing androconia. Journal of Chemical Ecology, 43, 843-857. http://doi:
638 10.1007/s10886-017-0867-3

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

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

645 Merrill, R. M., Rastas P., Martin S. H., Melo, M. C., Barker, S., Davey, J., ...Jiggins, C. D. (2019) Genetic
646 dissection of assortative mating behavior. PLOS Biology 17:e2005902. http://doi:
647 10.1371/journal.pbio.2005902

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

651 Moto, K., Suzuki, M. G., Hull, J. J., Kurata, R., Takahashi, S., Yamamoto, M., ...Matsumoto, S. (2004)
652 Involvement of a bifunctional fatty-acyl desaturase in the biosynthesis of the silkmoth, *Bombyx mori*, sex
653 pheromone. Proceedings of the National Academy of Sciences USA, 101, 8631-8636. http://doi:
654 10.1073/pnas.0402056101

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

657 Niehuis, O., Büllesbach, J., Gibson, J. D., Pothmann, D., Hanner, C., Mutti, N. S., ...Schmitt, T. (2013)
658 Behavioral and genetic analyses of *Nasonia* shed light on the evolution of sex pheromones. Nature, 494,
659 345-348. http://doi: 10.1038/nature11838

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

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

665 Picelli S., Björklund, Å. K., Reinius, B., Sagasser, S., Winberg, G., & Sandberg, R. (2014) Tn5 transposase
666 and tagmentation procedures for massively scaled sequencing projects. Genome Research, 24, 2033-
667 2040. http://doi: 10.1101/gr.177881.114

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

670 Reed, R. D., Papa, R., Martin, A., Hines, H. M., Counterman, B. A., Pardo-Diaz, C., ...McMillan, W. O.
671 (2011) *optix* Drives the Repeated Convergent Evolution of Butterfly Wing Pattern Mimicry. Science, 333,
672 1137-1141. http://doi: 10.1126/science.1208227

673 Riffell, J. A., Lei, H., Christensen, T. A., & Hildebrand, J. G. (2009) Characterization and Coding of
674 Behaviorally Significant Odor Mixtures. Current Biology, 19, 335-340. http://doi:
675 10.1016/j.cub.2009.01.041

676 Roelofs, W. L., Liu, W., Hao, G., Jiao, H., Rooney, A. P., & Linn, C. E., Jr. (2002) Evolution of moth sex
677 pheromones via ancestral genes. Proceedings of the National Academy of Sciences USA, 99, 13621-
678 13626. http://doi: 10.1073/pnas.152445399

679 Sakai, R., Fukuzawa, M., Nakano, R., Tatsuki, S., & Ishikawa, Y. (2009) Alternative suppression of
680 transcription from two desaturase genes is the key for species-specific sex pheromone biosynthesis in
681 two *Ostrinia* moths. Insect Biochemistry and Molecular Biology, 39, 62-67. http://doi:
682 10.1016/j.ibmb.2008.10.001

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

685 Sheck, A. L., Groot, A. T., Ward, C. M., Gemeno, C., Wang, J., Brownie, C., ...Gould, F. (2006) Genetics of
686 sex pheromone blend differences between *Heliothis virescens* and *Heliothis subflexa*: a chromosome
687 mapping approach. Journal of Evolutionary Biology, 19, 600-617. http://doi: 10.1111/j.1420-
688 9101.2005.00999.x

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

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

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

696 Van Belleghem, S. M., Rastas P., Papanicolaou, A., Martin, S. H., Arias, C. F., Supple, M. A., & Papa, R.
697 (2017) Complex modular architecture around a simple toolkit of wing pattern genes. Nature Ecology and
698 Evolution, 1, 0052. http://doi: 10.1038/s41559-016-0052

699 Via, S. (2012) Divergence hitchhiking and the spread of genomic isolation during ecological speciation-
700 with-gene-flow. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 367,
701 451-460. http://doi: 10.1098/rstb.2011.0260

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

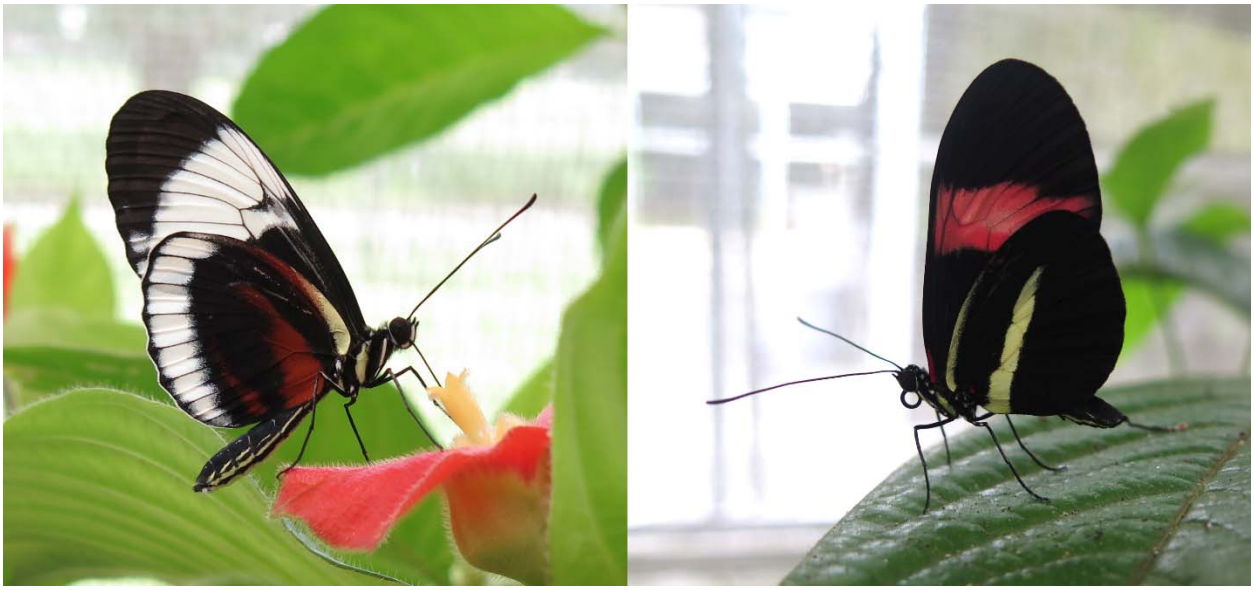
705 Wiley, C., Ellison, C. K., & Shaw, K. L. (2011) Widespread genetic linkage of mating signals and
706 preferences in the Hawaiian cricket *Laupala*. Proceedings of the Royal Society of London B: Biological
707 Sciences, 279, 1203–1209. http://doi: 10.1098/rspb.2011.1740

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

711 Zelle, K. M., Vernier, C., Leitner, N., Liang, X., Halloran, S., Millar, J. G., & Ben-Shahar, Y. (2020)
712 Chemoreceptor pleiotropy facilitates the functional coupling of mating pheromone synthesis and
713 perception. bioRxiv, 124305. http://doi: 10.1101/124305

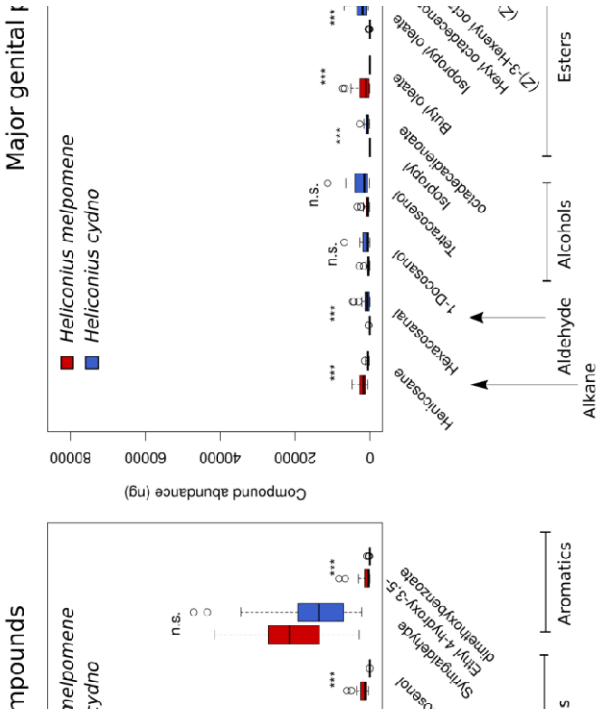
714 **Figures**

715 Figure 1: *Heliconius cydno chioneus* (left) and *Heliconius melpomene rosina* (right), the two butterfly
716 species whose putative pheromone components and wing androconial area are mapped in this study.

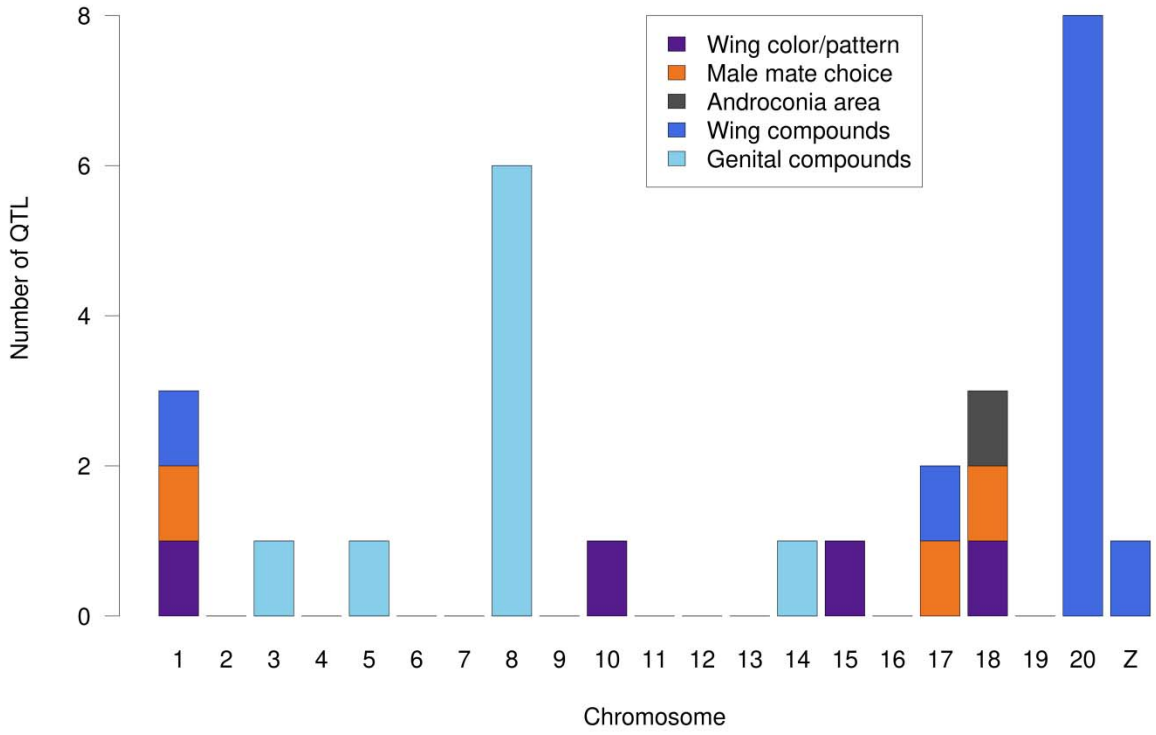


717

Figure 2: Major compounds found in the wings (left) and genitals (right) of *Heliconius melpomene* and *H. cydno*, arranged by compound type. Only those compounds comprising at least 1% of either species' pheromone bouquet are included. n.s., not significant; ***: $p < 0.005$. Boxplots: line is the median, box outline the first and third quartiles, whiskers the most extreme point no more than 1.5 times the interquartile range. [image rotated for reviewer ease of viewing]



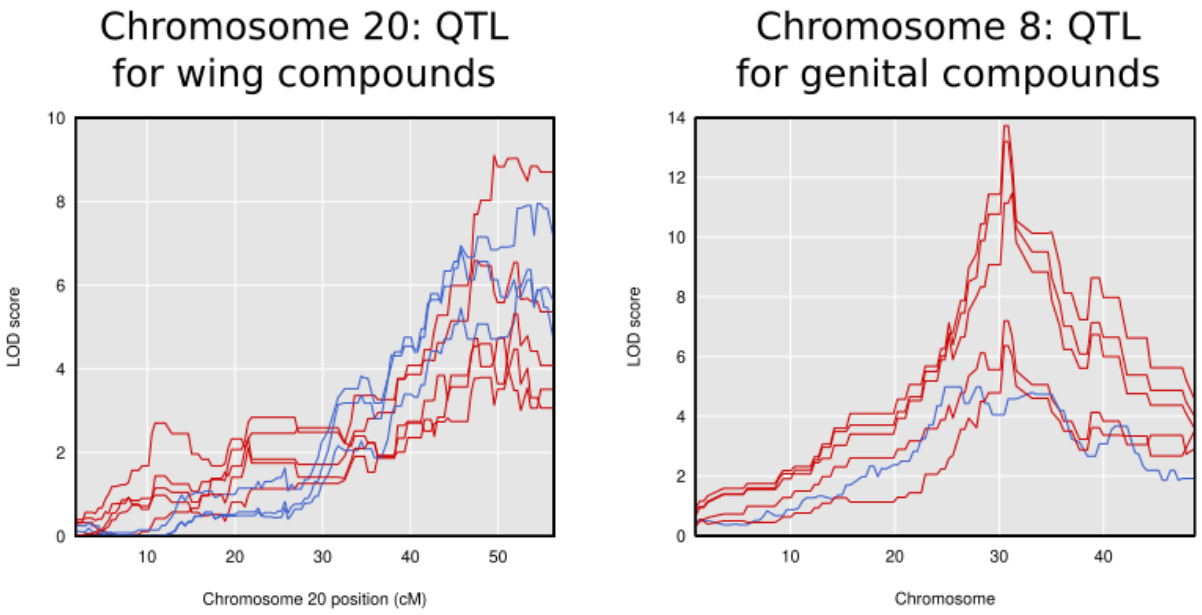
725 Figure 3: QTL for various traits across the *Heliconius melpomene* and *H. cydno* genomes. Wing
726 color/pattern loci represent the four major color loci (Jiggins, 2017); male mate choice loci are from
727 (Merrill et al., 2019); (relative) androconia area and Bonferroni-significant wing and genital compound
728 loci from this study.



729

730

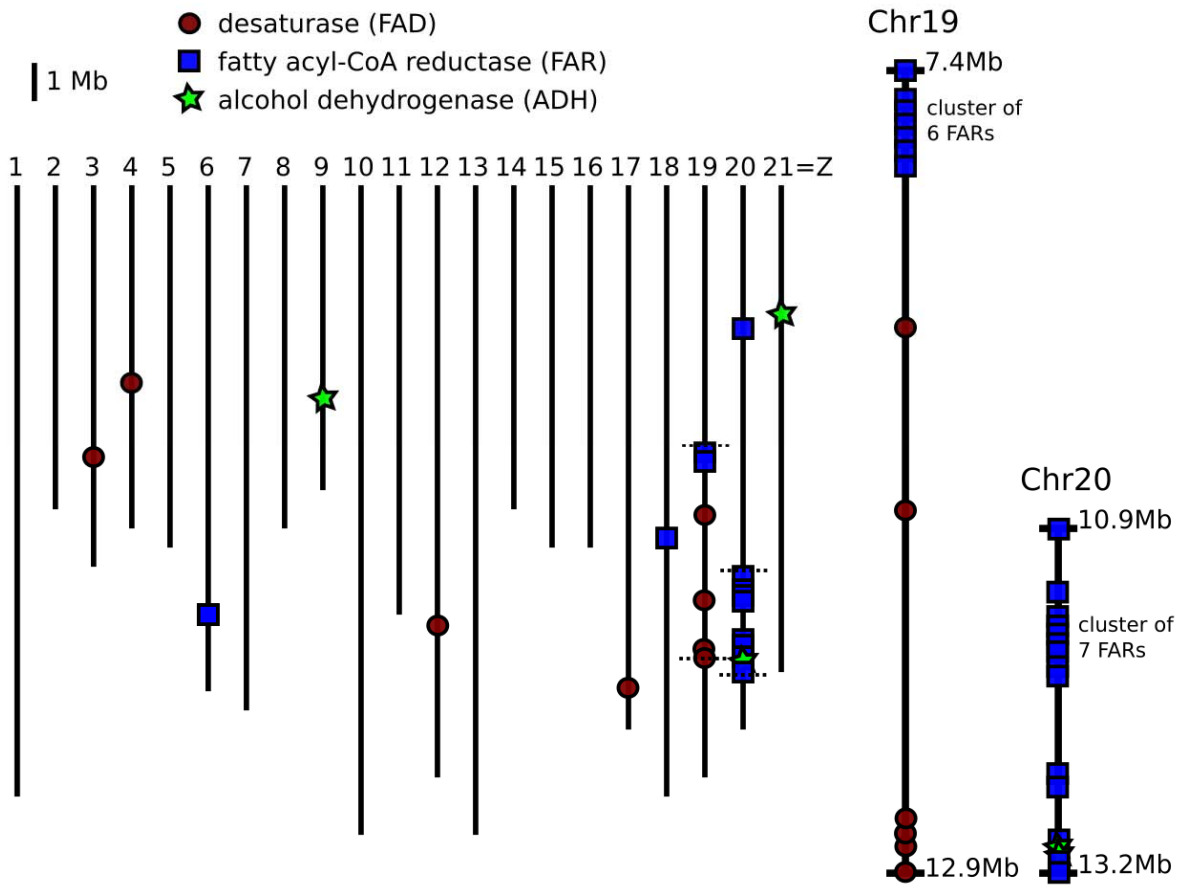
731 Figure 4: Chromosome-specific QTL plots for compounds with significant QTL (after Bonferroni
732 correction) on those chromosomes, with the entire chromosome length shown. Red: compound
733 mapped in backcrosses to *H. melpomene*; blue: compound mapped in backcrosses to *H. cydno*.



734

735

736 Figure 5: Location of candidate genes for fatty acid-derived compound biosynthesis across the genome
737 of *Heliconius melpomene*. Horizontal dotted lines on chromosomes 19 and 20 indicate the endpoints of
738 zoomed in regions.



739

740

741 **Tables**

742 Table 1: Full list of compounds found in the wings and genitalia of *Heliconius melpomene* and *H. cydno*,
743 details of statistical comparisons of amounts in both species, and cross direction(s) the compound was
744 mapped in, if any. Coefficient of variance in backcrosses is the variance within the backcross direction
745 divided by the mean amount found within that backcross direction, to control for variance differences
746 due to compound abundance differences.

747 **Wings**:

Compound	Compound type	Mean *H. cydno*	Mean *H. melpomene*	Kruskal-Wallis chi-squared statistic	E a v
Icosane	alkane	6.977	0.033	41.39738	
Henicosane	alkane	1152.641	12.447	38.66667	
Docosane	alkane	13.548	0.012	42.33465	
Tricosane	alkane	62.532	1.146	38.93777	
Pentacosane	alkane	4.568	8.458	3.73257	
Hexacosane	alkane	0.633	8.09	17.34831	
Octadecanol	saturated alcohol	absent	345.636	42.61856	
Methyloctadecanol	saturated alcohol	absent	12.232	42.61856	
Octadecanal	saturated aldehyde	0.569	783.420	39.55981	
Nonadecanal, methyl branched	saturated aldehyde	absent	4.504	24.23834	
Icosanal	saturated aldehyde	0.432	22.596	40.29682	
Tricosene (RI=2265)	alkene	3.112	absent	26.04236	
Tricosene (RI=2276)	alkene	0.405	7.215	35.29615	
(*Z*)-16-methyl-9-octadecenol	unsaturated alcohol	absent	7.501	40.26682	
(*Z*)-11-icosenol	unsaturated alcohol	28.174	772.157	38.66667	
(*Z*)-13-docosenol	unsaturated alcohol	0.037	54.766	42.10219	
(*Z*)-9-octadecenal	unsaturated aldehyde	absent	10.938	40.26682	
(*Z*)-11-icosenal	unsaturated aldehyde	3.29	150.687	39.75659	
(*Z*)-13-docosenal	unsaturated aldehyde	0.276	62.556	41.22627	
Naphthalene	aromatic	0.752	1.690	9.98090	

Syringaldehyde	aromatic	463.585	644.064	3.72414
3,5-Dimethoxy-4-hydroxy benzyl alcohol	aromatic	3.58	4.162	0.37288
Methyl salicylate	aromatic	2.848	3.348	0.34675
Ethyl 4-hydroxy-3,5-dimethoxybenzoate	aromatic	2.856	35.573	15.73593
Ethanone, 1,4-hydroxy-3,5-dimethoxyphenyl-	aromatic	1.992	2.981	2.91796
Propenone, 1,4-hydroxy-3,5-dimethoxyphenyl-	aromatic	2.052	3.932	5.66494
Benzyl cyanide	aromatic (nitrogenous)	1.813	absent	30.54161
Methyl 3-indoleacetate	aromatic (nitrogenous)	0.013	3.782	32.66847
Methyl 3-indolecarboxylate	aromatic (nitrogenous)	absent	2.098	25.99351
cis-Linalool oxide	terpenoid	3.407	absent	30.54161
Dihydroactinidiolide	terpenoid	3.020	1.236	11.50724

Genitals:

Compound	**Compound type**	**Mean *H. cydno***
Icosane	alkane	3.115
Henicosane	alkane	550.519
Docosane	alkane	absent
Tricosane	alkane	45.678
Pentacosane	alkane	8.393
11-Methylpentacosane	alkane	absent
Hexacosane	alkane	absent
Heptacosane	alkane	352.867
1-Docosanol	saturated alcohol	1187.774
1-Tetracosanol	saturated alcohol	171.351

1,3-Docosanediol (cyclic dimethylsilyl derivative)	saturated alcohol	297.605
1,3-Tricosanediol (cyclic dimethylsilyl derivative)	saturated alcohol	195.132
Hexacosanal	saturated aldehyde	1093.627
Hexyl 3-methylbutyrate	saturated ester	421.064
Hexyl hexanoate	saturated ester	111.474
Butyl hexadecanoate	saturated ester	absent
Octadecanoic acid ester	saturated ester	43.400
Butyl octadecanoate	saturated ester	absent
Isopropyl octadecanoate	saturated ester	97.088
Hexyl octadecanoate (RI=2550)	saturated ester	24.217
Hexyl octadecanoate (RI=2590)	saturated ester	415.139
Henicosene (RI=2068)	alkene	absent
Henicosene (RI=2072)	alkene	81.486
Tricosene (RI=2275)	alkene	217.571
Tricosene (RI=2270)	alkene	64.338
(*Z*)-13-Docosen-1-ol	unsaturated alcohol	71.558
Tetracosenol	unsaturated alcohol	2559.806
(*Z*)-3-Hexenyl isobutyrate	unsaturated ester	26.306
Hexyl hexadecenoate	unsaturated ester	63.084
Butyl oleate	unsaturated ester	absent
Isopropyl oleate	unsaturated ester	2259.264
(*Z*)-3-Hexenyl hexadecanoate	unsaturated ester	97.210
Isopropyl octadecadienoate	unsaturated ester	740.075
Isopropyl 9,12-octadecadienoate	unsaturated ester	405.054
Hexyl octadecenoate and (Z)-3-Hexenyl octadecenoate	unsaturated ester	2898.234
Hexenyl octadecatrienoate & (*Z*)-3-Hexenyl octadecatrienoate	unsaturated ester	367.972
(*Z*)-3-Hexenyl octadecanoate	unsaturated ester	35.961
5-Decanolide	lactone	36.466
15-Tetradecanolide	macrolide	41.895
Brassicalactone	macrolide	127.495
16-Hexadecanolide	macrolide	89.514
Hexadecenolide	macrolide	165.488

Hexadecen-11-olide	macrolide	110.815
9,11-Hexadecadien-11-olide	macrolide	72.479
12-Octadecanolide	macrolide	105.430
18-Octadecanolide	macrolide	168.048
Octadecen-11-olide	macrolide	175.211
(*E*)-Octadec-9-en-12-olide	macrolide	345.939
(*Z*)-9-Octadecen-11-olide	macrolide	7508.774
(*Z*)-9-Octadecen-13-olide	macrolide	5855.658
9-Octadecen-18-olide	macrolide	350.327
(*Z*9,*E*11)- 9,11-Octadecadien -13-olide	macrolide	19714.248
Octadeca-9,11-dien-13-olide and 11-Octadecanolide	macrolide	77.716
Octadecadienolide (RI=2095)	macrolide	642.378
Octadecadienolide (RI=2178)	macrolide	101.588
(*Z*9,*E*11,*Z*15)-9,11,15- Octadecatrien-13-olide	macrolide	1052.733
13-Icosanolide	macrolide	45.901
Icosen-13-olide	macrolide	781.655
Docosen-22-olide	macrolide	109.681
Tetracosenolide	macrolide	489.085
Benzyl cyanide	aromatic (nitrogenous)	287.978
2-*sec*-Butyl-3-methoxypyrazine	aromatic (nitrogenous)	3.569
2-Isobutyl-3-methoxypyrazine	aromatic (nitrogenous)	21.148
β-Myrcene	terpenoid	absent
(*Z*)-β-Ocimene	terpenoid	0.930
(*E*)-β-Ocimene	terpenoid	absent
Alloocimene	terpenoid	absent
Dihydroedulan II	terpenoid	20.336

Table 2: Details of compounds with Bonferroni-significant QTL in *Heliconius* wing androconia. a: from (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)**	**Percent parental difference explained by genotype at QTL position**

(*Z*)-9-Octadecenal	Unsaturated aldehyde	10.94	Absent	1[m]	9955174[m]	33.544%[m]
Benzyl cyanide	Aromatic (nitrogenous)	Absent	1.813	17[m]	5596551[m]	67.126%[m]
Octadecanal[a]	Saturated aldehyde	783.4	0.569	20[c]; 20[m]	10947362[c]; 11725848[m]	0.602%[c]; 41.411%[m]
Henicosane	Alkane	12.45	1152	20[m]	12126129[m]	11.291%[m]
Methyloctadecanol	Saturated alcohol	12.23	Absent	20[m]	12374765[m]	43.435%[m]
1-Octadecanol[a]	Saturated alcohol	345.6	Absent	20[m]	12652775[m]	27.709%[m]
(*Z*)-16-Methyl-9-octadecenol	Unsaturated alcohol	7.501	Absent	20[m]	12652775[m]	38.888%[m]
(*Z*)-13-Docosenal	Unsaturated aldehyde	62.56	0.276	20[c]	12917911[c]	28.582%[c]
(*Z*)-11-Icosenal	Unsaturated aldehyde	150.7	3.290	20[c]	13267611[c]	22.717%[c]
Methyl 3-indoleacetate	Aromatic (nitrogenous)	3.782	0.013	Z[c]	8478732[c]	59.220%[c]

Table 3: Details of compounds with Bonferroni-significant QTL in *Heliconius* genitals. a: from (Darragh 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)**	**Bonferroni-significant QTL on chromosome**	**QTL position on chromosome (bp)**	**Percent parental difference explained by genotype at QTL position**
Hexyl 3-methylbutyrate	Saturated ester	Absent	421.1	3[m]	234227[m]	5.863%[m]
2-*sec*-Butyl-3-methoxypyrazine	Aromatic	32.66	3.569	5[c]	4427722[c]	38.866%[c]

	(nitrogenous)					
12-Octadecanolide	Macrolide	Absent	105.4	8[c]	4512378[c]	31.434%[c]
(*Z*)-9-Octadecen-11-olide	Macrolide	Absent	7509	8[m]	5341821[m]	12.192%[m]
(*Z*9,*E*11)- 9,11-Octadecadien -13-olide	Macrolide	Absent	19714	8[m]	5341821[m]	2.826%[m]
Hexadecenolide	Macrolide	6.866	165.5	8[m]	5341821[m]	170.403%[m]
Hexyl octadecanoate	Saturated ester	121.8	24.21	8[m]	5987924[m]	46.335%[m]
(*Z*)-9-Octadecen-13-olide	Macrolide	7.282	5856	8[m]	6072562[m]	15.975%[m]
Isopropyl oleate	Unsaturated ester	50.22	2259	14[m]	7370836[m]	16.385%[m]

Table 4: Details of quantitative trait loci passing a significance threshold of either p = 0.05 or a Bonferroni-corrected threshold.

Wing compounds			
Compound (highlighted = Bonferroni-significant QTL)	**Compound type**	**Mapped in which backcross direction?**	**Chromosc**
Henicosane	alkane	to *H. melpomene*	
Icosane	alkane	to *H. melpomene*	
Methyloctadecanol	saturated alcohol	to *H. melpomene*	
1-Octadecanol	saturated alcohol	to *H. melpomene*	
Icosanal	saturated aldehyde	to *H. cydno*	
Octadecanal	saturated aldehyde	to *H. melpomene*	
Octadecanal	saturated aldehyde	to *H. cydno*	
(*Z*)-11-Icosenol	unsaturated alcohol	to *H. cydno*	
(*Z*)-13-Docosenol	unsaturated alcohol	to *H. melpomene*	
(*Z*)-16-Methyl-9-octadecenol	unsaturated alcohol	to *H. melpomene*	
(*Z*)-11-Icosenal	unsaturated aldehyde	to *H. cydno*	

(*Z*)-13-Docosenal	unsaturated aldehyde	to *H. cydno*
(*Z*)-9-Octadecenal	unsaturated aldehyde	to *H. melpomene*
(*Z*)-9-Octadecenal	unsaturated aldehyde	to *H. melpomene*
Benzyl cyanide	aromatic (nitrogenous)	to *H. melpomene*
Benzyl cyanide	aromatic (nitrogenous)	to *H. melpomene*
Benzyl cyanide	aromatic (nitrogenous)	to *H. cydno*
Methyl 3-indoleacetate	aromatic (nitrogenous)	to *H. cydno*

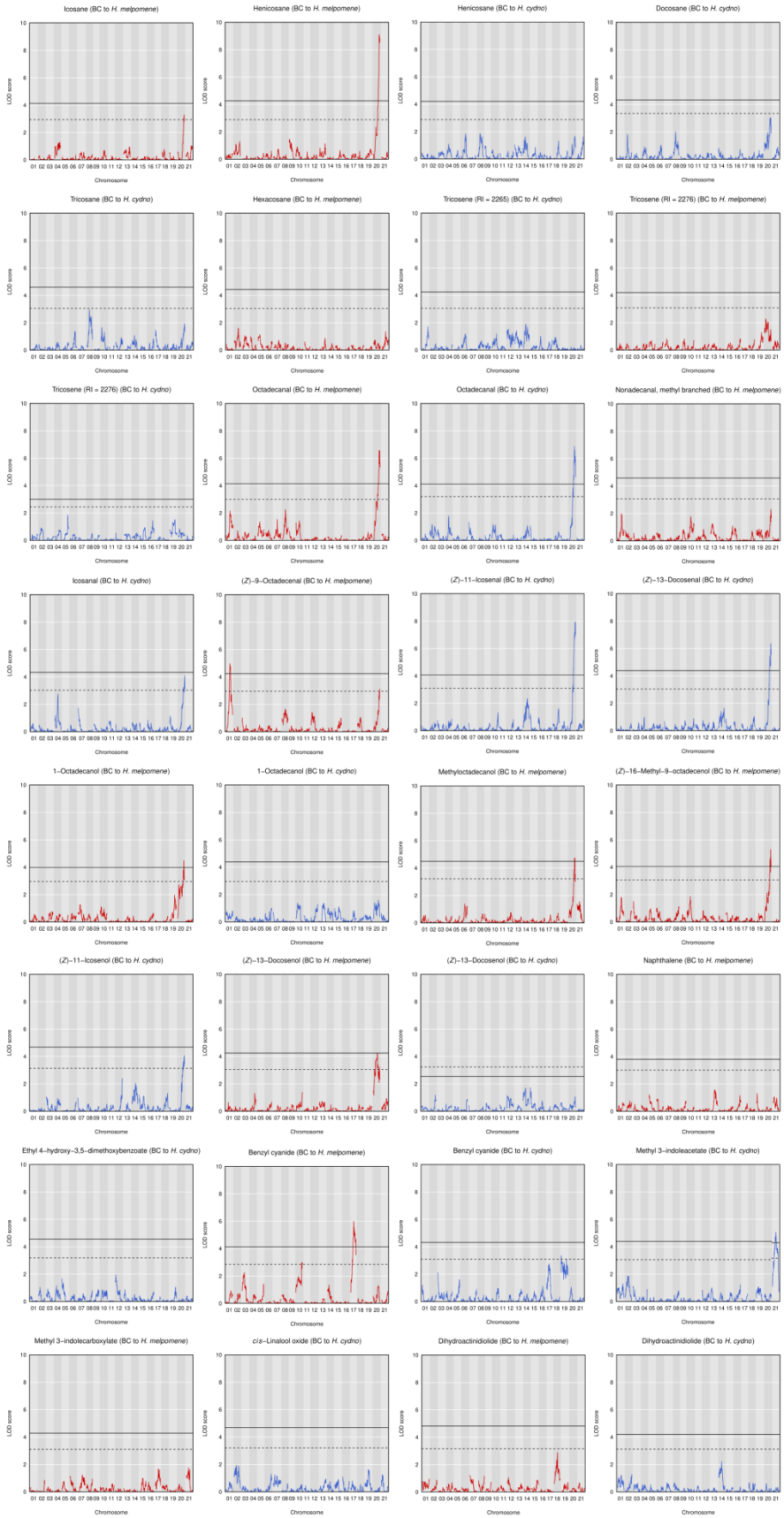
Genital compounds

Compound (highlighted = Bonferroni-significant QTL)	**Compound type**	**Mapped in which backcross direction?**	**Chromosc**
Docosane	alkane	to *H. melpomene*	
Henicosane	alkane	to *H. cydno*	
Butyl hexadecanoate	saturated ester	to *H. melpomene*	
Hexyl-3-methylbutyrate	saturated ester	to *H. melpomene*	
Hexyloctadecanoate (RI=2590)	saturated ester	to *H. melpomene*	
Octadecanoic acid ester	saturated ester	to *H. cydno*	
Butyl oleate	unsaturated ester	to *H. melpomene*	
Butyl oleate	unsaturated ester	to *H. melpomene*	
Butyl oleate	unsaturated ester	to *H. cydno*	
Isopropyl oleate	unsaturated ester	to *H. melpomene*	
2-*sec*-Butyl-3-methoxypyrazine	aromatic (nitrogenous)	to *H. cydno*	
Benzyl cyanide	aromatic (nitrogenous)	to *H. melpomene*	
(*Z*)-9-Octadecen-11-olide	macrolide	to *H. melpomene*	
(*Z*)-9-Octadecen-13-olide	macrolide	to *H. melpomene*	
(*Z*9,*E*11)- 9,11-Octadecadien -13-olide	macrolide	to *H. melpomene*	
12-Octadecanolide	macrolide	to *H. cydno*	
9-Octadecen-18-olide	macrolide	to *H. cydno*	
Brassicalactone	macrolide	to *H. cydno*	
Icosen-13-olide	macrolide	to *H. melpomene*	
Hexadecen-11-olide	macrolide	to *H. cydno*	
Hexadecenolide	macrolide	to *H. melpomene*	
(*Z*9,*E*11,*Z*15)-9,11,15- Octadecatrien-13-olide	macrolide	to *H. melpomene*	
(*E*)-β-Ocimene	terpenoid	to *H. cydno*	

759

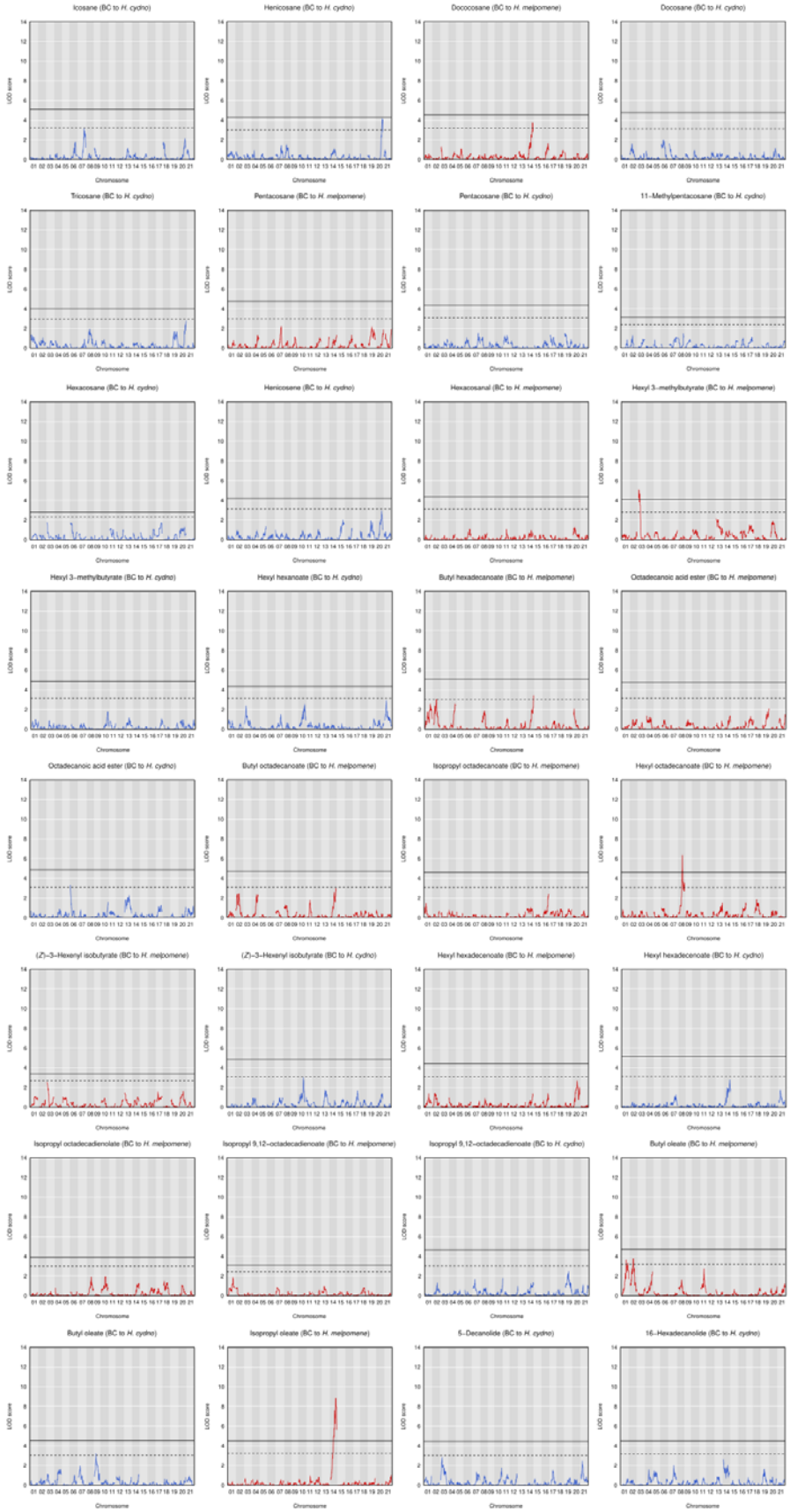
760 **Appendix: Supplemental figures**

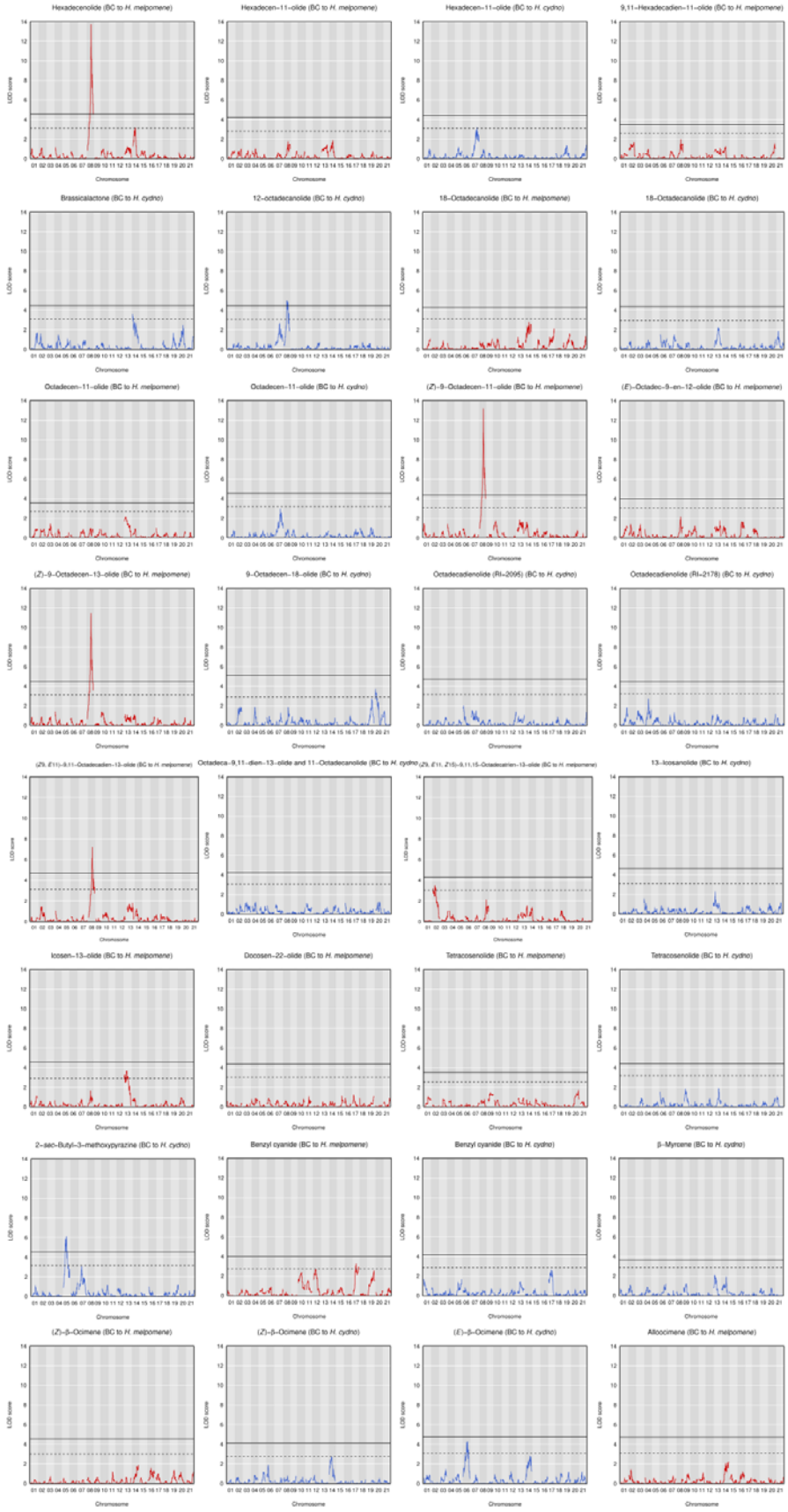
761 Supplemental Figure 1: QTL plots for all mapped wing compounds. Dashed horizontal line: $p = 0.05$
762 significance threshold. Solid horizontal line: Bonferroni-corrected significance threshold. BC, backcross.


Icosane (BC to H. melpomene)
Henicosane (BC to H. melpomene)
Henicosane (BC to H. cydno)
Docosane (BC to H. cydno)
Tricosane (BC to H. cydno)
Hexacosane (BC to H. melpomene)
Tricosene (RI = 2265) (BC to H. cydno)
Tricosene (RI = 2276) (BC to H. melpomene)
Tricosene (RI = 2276) (BC to H. cydno)
Octadecanal (BC to H. melpomene)
Octadecanal (BC to H. cydno)
Nonadecanal, methyl branched (BC to H. melpomene)
Icosanal (BC to H. cydno)
(Z)−9−Octadecenal (BC to H. melpomene)
(Z)−11−Icosenal (BC to H. cydno)
(Z)−13−Docosenal (BC to H. cydno)
1−Octadecanol (BC to H. melpomene)
1−Octadecanol (BC to H. cydno)
Methyloctadecanol (BC to H. melpomene)
(Z)−16−Methyl−9−octadecenol (BC to H. melpomene)
(Z)−11−Icosenol (BC to H. cydno)
(Z)−13−Docosenol (BC to H. melpomene)
(Z)−13−Docosenol (BC to H. cydno)
Naphthalene (BC to H. melpomene)
Ethyl 4−hydroxy−3,5−dimethoxybenzoate (BC to H. cydno)
Benzyl cyanide (BC to H. melpomene)
Benzyl cyanide (BC to H. cydno)
Methyl 3−indoleacetate (BC to H. cydno)
Methyl 3−indolecarboxylate (BC to H. melpomene)
cis−Linalool oxide (BC to H. cydno)
Dihydroactinidiolide (BC to H. melpomene)
Dihydroactinidiolide (BC to H. cydno)
LOD score
Chromosome

763

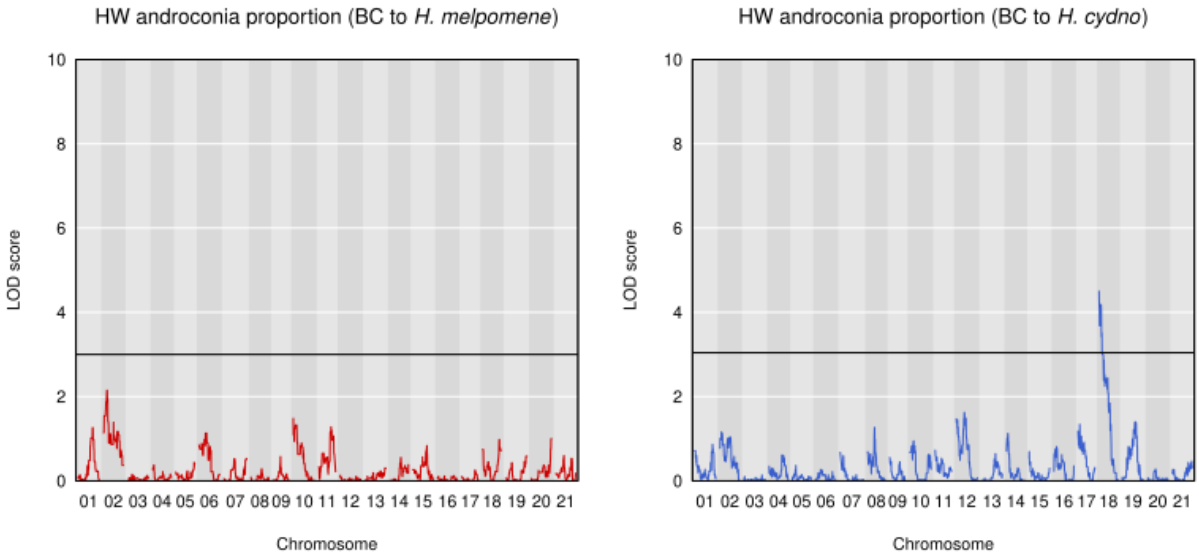
764 Supplemental Figure 2 (in two parts): QTL plots for all mapped genital compounds. Dashed horizontal
765 line: p = 0.05 significance threshold. Solid horizontal line: Bonferroni-corrected significance threshold.
766 BC, backcross.


Icosane (BC to H. cydno)
Henicosane (BC to H. cydno)
Dococosane (BC to H. melpomene)
Docosane (BC to H. cydno)
Tricosane (BC to H. cydno)
Pentacosane (BC to H. melpomene)
Pentacosane (BC to H. cydno)
11-Methylpentacosane (BC to H. cydno)
Hexacosane (BC to H. cydno)
Henicosane (BC to H. cydno)
Hexacosanal (BC to H. melpomene)
Hexyl 3-methylbutyrate (BC to H. melpomene)
Hexyl 3-methylbutyrate (BC to H. cydno)
Hexyl hexanoate (BC to H. cydno)
Butyl hexadecanoate (BC to H. melpomene)
Octadecanoic acid ester (BC to H. melpomene)
Octadecanoic acid ester (BC to H. cydno)
Butyl octadecanoate (BC to H. melpomene)
Isopropyl octadecanoate (BC to H. melpomene)
Hexyl octadecanoate (BC to H. melpomene)
(Z)-3-Hexenyl isobutyrate (BC to H. melpomene)
(Z)-3-Hexenyl isobutyrate (BC to H. cydno)
Hexyl hexadecenoate (BC to H. melpomene)
Hexyl hexadecenoate (BC to H. cydno)
Isopropyl octadecadienolate (BC to H. melpomene)
Isopropyl 9,12-octadecadienoate (BC to H. melpomene)
Isopropyl 9,12-octadecadienoate (BC to H. cydno)
Butyl oleate (BC to H. melpomene)
Butyl oleate (BC to H. cydno)
Isopropyl oleate (BC to H. melpomene)
5-Decanolide (BC to H. cydno)
16-Hexadecanolide (BC to H. cydno)
LOD score
Chromosome


Hexadecenolide (BC to H. melpomene)
Hexadecen-11-olide (BC to H. melpomene)
Hexadecen-11-olide (BC to H. cydno)
9,11-Hexadecadien-11-olide (BC to H. melpomene)
Brassicalactone (BC to H. cydno)
12-octadecanolide (BC to H. cydno)
18-Octadecanolide (BC to H. melpomene)
18-Octadecanolide (BC to H. cydno)
Octadecen-11-olide (BC to H. melpomene)
Octadecen-11-olide (BC to H. cydno)
(Z)-9-Octadecen-11-olide (BC to H. melpomene)
(E)-Octadec-9-en-12-olide (BC to H. melpomene)
(Z)-9-Octadecen-13-olide (BC to H. melpomene)
9-Octadecen-18-olide (BC to H. cydno)
Octadecadienolide (RI-2095) (BC to H. cydno)
Octadecadienolide (RI-2178) (BC to H. cydno)
(Z9, E11)-9,11-Octadecadien-13-olide (BC to H. melpomene)
Octadeca-9,11-dien-13-olide and 11-Octadecanolide (BC to H. cydno)
(Z9, E11, Z15)-9,11,15-Octadecatrien-13-olide (BC to H. melpomene)
13-Icosanolide (BC to H. cydno)
Icosen-13-olide (BC to H. melpomene)
Docosen-22-olide (BC to H. melpomene)
Tetracosenolide (BC to H. melpomene)
Tetracosenolide (BC to H. cydno)
2-sec-Butyl-3-methoxypyrazine (BC to H. cydno)
Benzyl cyanide (BC to H. melpomene)
Benzyl cyanide (BC to H. cydno)
β-Myrcene (BC to H. cydno)
(Z)-β-Ocimene (BC to H. melpomene)
(Z)-β-Ocimene (BC to H. cydno)
(E)-β-Ocimene (BC to H. cydno)
Alloocimene (BC to H. melpomene)
LOD score
Chromosome

769 Supplemental Figure 3: QTL plots for hindwing androconial proportion. Solid horizontal line: p = 0.05

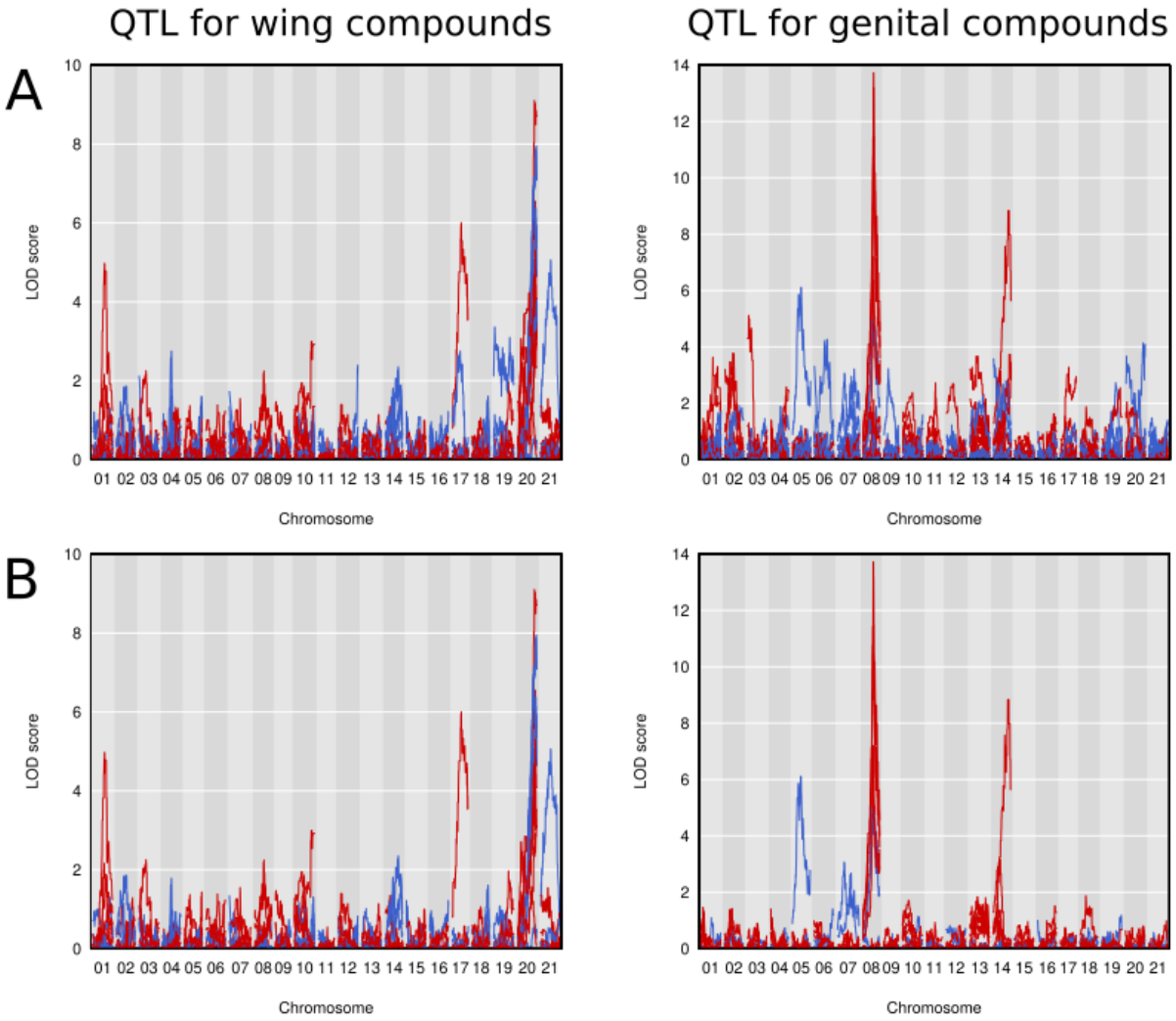
770 significance threshold. HW, hindwing. BC, backcross.



771

772

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



777

778

779