A major locus controls a biologically active pheromone component in *Heliconius melpomene*

Kelsey J.R.P. Byers[1,2,9], Kathy Darragh[1,2,9], Jamie Musgrove[2], Diana Abondano Almeida[2,3], Sylvia Fernanda Garza[2,4], Ian A. Warren[1], Pasi Rastas[5], Marek Kucka[6], Yingguang Frank Chan[6], Richard M. Merrill[7], Stefan Schulz[8], W. Owen McMillan[2], Chris D. Jiggins[1,2,10]

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

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

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

[4] *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*

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

[6] *Friedrich Miescher Laboratory of the Max Planck Society, Tuebingen, Germany*

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

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

[9] *These authors contributed equally to this work*

[10] *To whom correspondence should be addressed: c.jiggins@zoo.cam.ac.uk*

Running title: Genetics of bioactive pheromones in *Heliconius*

Abstract

Pheromones are important for courtship and mate choice in many insects, but we know relatively little of their role in butterflies. The butterfly *Heliconius melpomene* uses a complex blend of wing androconial compounds during courtship. Electroantennography in *H. melpomene* and its close relative *H. cydno* showed that responses to androconial extracts were not species-specific. Females of both species responded more strongly to the *H. cydno* extract, suggesting conservation of peripheral nervous system elements across the two species. Individual blend components provoked little to no response, with the exception of octadecanal, a major component of the *H. melpomene* blend. Supplementing octadecanal on the wings of octadecanal-rich *H. melpomene* males led to an increase in the time until mating, demonstrating the bioactivity of octadecanal in *Heliconius*. Using quantitative trait locus (QTL) mapping, we identified a single locus on chromosome 20 responsible for 41% of the parental species' difference in octadecanal production. This QTL does not overlap with any of the major wing color or mate choice loci, nor does it overlap with known regions of elevated or reduced F_{ST}. A set of 16 candidate fatty acid biosynthesis genes lies underneath the QTL.

Introduction

Chemical communication is the oldest form of sensory communication, and is found across the tree of life from bacteria (Taga & Bassler 2003) to humans (Jacob et al 2002; Wysocki & Preti 2004). In Lepidoptera, work on pheromones has largely focused on female pheromones of nocturnal moths, where relatively simple alterations to pheromones can produce drastic differences in mate attraction. For example, different populations of the corn borer moth *Ostrinia nubialis* exhibit a simple *cis* to *trans* switch in the pheromone 11-tetradecenyl acetate (Kochansky et al 1975) that leads to partial reproductive isolation between them (Lassance et al 2010). Sex pheromones have been less well studied in day-flying butterflies, where visual signaling is often assumed to play a more dominant role in mate choice (Vane-Wright & Boppré 1993). However, male butterflies also emit close-range pheromone bouquets which may act in concert with other wing pattern and behavioral cues (Mérot et al 2015), and are important in mate choice (Darragh et al 2017).

Despite the potential importance of pheromones in butterflies, pheromones have so far been identified in only a handful of butterfly species (Meinwald 1969, Pliske & Eisner 1969, Grula et al 1980, Nishida et al 1996, Schulz & Nishida 1996, Andersson 2007, Nieberding et al 2008, Yildizhan et al 2009). Within

butterflies, the genetic basis of pheromone production is known only in *Bicyclus anynana* (Lienard et al 2014). As in other diurnal butterflies, the chemical bouquets of *Heliconius* (Darragh et al 2017, Mann et al 2018), are often complex, both in identity and quantity of compounds; however, just a few individual compounds may be biologically active pheromones. Identifying which components of these complex chemical bouquets are responsible for pheromonal communication is a considerable challenge. This is particularly true as even minor compounds can have major effects (McCormick et al 2014, Chen et al 2018). Determining pheromone bioactivity requires screening compounds via physiological activity followed by behavioral verification, and in many cases pheromone bouquet composition is known but the bioactive components remain unidentified.

To better characterize male butterfly pheromone production and the role it plays in mating and species recognition, we first comprehensively analyzed pheromone bouquets of the closely related *Heliconius melpomene* and *H. cydno* butterflies to determine the most important bioactive compounds. We further mapped a genetic locus likely responsible for the differential biosynthesis of the main fatty-acid compound between the two species. The two species are closely related yet strongly reproductively isolated (Jiggins 2017). Over the past decade there has been considerable research into genomic architecture of differences in wing pattern and mate preference (Jiggins 2017; Merrill et al 2019). Surprisingly, both wing color and mating preferences between these species have a relatively simple genetic basis with a large proportion of the difference between parental forms being controlled by a small handful of loci of large effect (Jiggins 2017, Merrill et al 2019). This has important implications for speciation, as theory predicts that large effect loci should facilitate speciation in the face of gene flow (Via 2012). Similarly, clustering of genetic loci that contribute to speciation can also make speciation easier (Merrill et al 2010, Smadja & Butlin 2011), and there is evidence for tight linkage of a gene controlling wing pattern and a QTL underlying divergent preference behaviors (Merrill et al 2019). Here, we identify an electrophysiologically active compound, octadecanal, explore its behavioral role, and investigate the genetic basis of its production in *H. melpomene* using QTL mapping.

Results

Androconial chemical bouquets of *Heliconius melpomene* and *H. cydno*

In order to identify candidate pheromone compounds, we first investigated the distribution of chemical compounds on the wings of *H. cydno* for comparison with published data on *H. melpomene*. We found that the bouquet of *H. cydno* was simpler than that of *H. melpomene*, with 7 main compounds versus 21

in *H. melpomene*. *Heliconius cydno* also had a less abundant overall bouquet, with an individual total of 1787 ± 776 ng vs. 3174 ± 1040 ng in *H. melpomene* (Table S1). Most of the main compounds (defined as those occurring in at least 90% of individuals) in *H. cydno* were linear alkanes (4 of 7 compounds), while *H. melpomene* has a more diverse array of different compound classes. Of the five major compounds (> 100ng per individual) in *H. melpomene*, only syringaldehyde was found in similar amounts in *H. cydno*; the major compounds octadecanal, 1-octadecanol, (*Z*)-11-icosenal, and (*Z*)-11-icosenol were absent or found in very low amounts in *H. cydno*. Comparison with previously published data for other *Heliconius* species, all of which lack octadecanal in the large amounts seen in *H. melpomene*, demonstrates that this high level of octadecanal is an evolutionarily derived state in *H. melpomene* (Mann et al 2017, Darragh et al 2019b). When focusing on the hindwing androconia of *H. cydno*, only two compounds [syringaldehyde (24.7% of the hindwing androconial bouquet) and (*Z*)-11-icosenol (1.7%)] were specific to this region. For details, please see SI text, Figures S1-S3, and Tables S1-S2.

Electroantennographic responses to con- and heterospecific pheromone stimulus sets

We next investigated the electroantennographic (EAG) response of both species to natural con- and heterospecific pheromone bouquets extracted from adult male butterflies. Females of both *H. melpomene* and *H. cydno* responded more strongly (i.e. showed a larger voltage displacement from the antennal baseline) to both natural wing extracts (Mnat and Cnat, respectively) as compared with the control solvent (dichloromethane plus 1 ng/µL 2-tetradecyl acetate, hereafter "DCM+IS") (Figure 1, Figure S4; see Table S3 for statistical details). Males of both species also responded to both wing extracts. Females of both species showed stronger responses to *H. cydno* wing extracts than to *H. melpomene* wing extracts. Responses differed in males, with *H. melpomene* males responding equally to both species' wing extracts, while males of *H. cydno* responded more strongly to wing extracts of *H. melpomene*.

We then explored antennal responses to synthetic compound blends. These were based on the most abundant compounds from each species (see Methods). Responses to the synthetic *H. melpomene* wing blend were mixed in *H. melpomene* females. In the first stimulus set (comparison of hetero- and conspecific stimuli) the synthetic blend (Msyn) yielded stronger responses than the natural *H. melpomene* wing extracts (Figure 1, Table S3) but in the second stimulus set (individual pheromone components) the synthetic blend yielded weaker responses than the natural *H. melpomene* wing extracts (Figure 2, Table S3). A similar pattern was seen in male *H. melpomene*. Both sexes of *H. cydno*

111 produced only a modest response to the synthetic *H. cydno* wing blend (Csyn) when compared with the
112 hexane solvent. In all cases this response was lower than to natural *H. cydno* wing extract.

113 Electroantennographic responses to individual pheromone components from both species

114 Finally, we explored the responses to individual compounds to identify specific biologically active
115 pheromone components. Only octadecanal differed significantly from the controls in all four species-sex
116 combinations (Figure 2, Table 1, Table S3). Both sexes of *H. cydno* also showed altered responses to the
117 shared compound (*Z*)-11-icosenol, with an increased response in *H. cydno* males and a decreased
118 response in *H. cydno* females. Both *H. cydno* females and *H. melpomene* males had decreased responses
119 to (*Z*)-11-icosenal. *Heliconius cydno* females also showed a decreased response to (*Z*)-13-docosenal.
120 Male *H. cydno* had increased responses to 1-octadecanol and henicosane. No other compound-species-
121 sex combinations were significantly different from hexane. The shared compound syringaldehyde was
122 the only pheromone component to show no change in EAG amplitude relative to the solvents in any
123 species-sex combination.

124 In females of both species, response to octadecanal was stronger than response to the conspecific
125 synthetic mixture. By contrast, male *H. melpomene* responded equally to the conspecific synthetic
126 mixture and octadecanal, while male *H. cydno* responded equally to the conspecific synthetic mixture
127 and both its components syringaldehyde and henicosane, with no evidence for a synergistic mixture
128 effect. Individual components (with the exception of octadecanal) were not equivalent to the synthetic
129 blends for females of both *H. melpomene* and *H. cydno*, suggesting a synergistic mixture effect. Antennal
130 responses to a given stimulus can change over time, and this may reflect biological processes of
131 neuronal adaptation. Female *H. melpomene* adapted equally quickly to octadecanal and the natural and
132 synthetic *H. melpomene* pheromones, while adaptation to other stimuli was equal to the control, further
133 supporting octadecanal's salience as the main pheromone in *H. melpomen*e (see SI text, Figures S5-S6,
134 and Table S4).

135 Behavioral response to octadecanal supplementation in *H. melpomene*

136 We next confirmed a behavioral response to the most physiologically active substance, octadecanal, one
137 of the dominant compounds in *H. melpomene*. A total of 29 behavioral experiments were conducted,
138 with mating observed in 18 (62%); one trial was excluded due to wing damage, leaving 17 successful
139 matings. With our small sample size, we found no evidence that females showed a preference for either
140 treatment, mating with the hexane male 11 times (65%) and with the octadecanal male the remaining 6

times (35%) ($p = 0.332$). However, mating latency (time from experiment onset to mating) was significantly longer for the octadecanal matings (average 88.5 minutes) than for the hexane matings (average 43.7 minutes; $t = 2.7848$, $df = 8.2491$, $p = 0.023$). There was no evidence that this mating latency was due to evaporation of the octadecanal treatment, as there was no detectable drop in octadecanal quantity in the hindwing androconia over time [comparison of 30 minute and 2 hour treatments (Figure S7)]. Furthermore, little octadecanal rubbed off onto the forewing overlap region. Interestingly, although about 5.5x as much octadecanal as normal should have been present on the wings of treated males, only about 2-2.5x was seen after 30 minutes (the time at which the female would be introduced to the two males), suggesting some of the added octadecanal was lost before the start of the behavioral experiments, perhaps due to oxidization or pheromone hydrolysis, as has been shown in some moths (Ferkovich et al 1982).

Genetic basis of octadecanal production in *H. melpomene*

Analysis of octadecanal production by F_1 males showed that the *H.* cydno octadecanal phenotype (little to no octadecanal) is dominant over the octadecanal-rich *H. melpomene* phenotype, and octadecanal production segregates in backcrosses to *H. melpomene* (Figure S10). Using the variance within the backcross individuals, we calculated a Castle-Wright estimator of 0.81 loci, suggesting a potentially monogenic basis for octadecanal production in *H. melpomene*. Quantitative trait locus (QTL) mapping revealed a single significant peak on chromosome 20 (Figure 3a, Figure S11). The chromosome 20 peak remained significant when kinship was taken into account and explained 41.31% of the difference between the two parent species. Bayesian confidence intervals for the peak on chromosome 20 were identical with and without kinship, spanning a range of 46.9-56.37cM with the peak at 47.66cM (with or without kinship), corresponding to a physical range of 3.4Mb along chromosome 20. To ensure that our findings were replicable across individual families, we also constructed effect plots at the kinship peak for each family separately, and all showed the same directionality (Figure 3b). The peak on chromosome 20 does not overlap with any of the major wing color loci (Jiggins 2017; van Belleghem et al 2017), nor does it overlap with mate choice QTL (Merrill et al 2019). The confidence interval region contains 160 genes, all of which represent potential candidates for octadecanal production.

We next evaluated the evidence in support of the 160 genes in this interval to identify top candidates for octadecanal production. In total, 14 were putative fatty-acyl CoA reductases (FARs), which catalyze the conversion of fatty acids to alcohols via a bound aldehyde intermediate (Table S5). Octadecanal is most likely produced via this pathway (Figure 3c), either as a direct product of a FAR-catalyzed

conversion of 18-carbon stearic acid (by releasing the bound intermediate directly) or as a product of a further dehydrogenation of the alcohol intermediate (octadecanol) to the aldehyde product. Two candidate alcohol dehydrogenases, which might catalyze this reaction, were also contained within the region, yielding a total of 16 candidates. To ascertain whether octadecanol might serve as the precursor to octadecanal in *H. melpomene*, we also searched for QTL underlying octadecanol production (Castle-Wright estimator of 0.71 loci), and found a very similar pattern to octadecanal, with a single QTL peak on chromosome 20 (Figure 3d) explaining 25.36% of the difference between the two parent species. This peak broadly overlapped the octadecanal peak, with a much broader confidence interval from 10.91-56.37cm (12.9Mb) and a peak at 51.82cM regardless of whether kinship was taken into account (Figure S11). The 14 FARs in the region are highly clustered, with a set of eight found within a 133kb region. Sequence comparison between the *H. melpomene* and *H. cydno* alleles showed that nearly all of these genes harbor nonsynonymous SNPs between the two species (Table S5, Supplementary Data 1). One gene showed no coding SNPs between the two species; nine had between three and 24 nonsynonymous SNPs; and four had more substantial changes, including missing exons and frameshifts. The final two genes (one alcohol dehydrogenase and one FAR) could not be found in the *H. cydno* genome, and may instead represent annotation or assembly errors in *H. melpomene* or, alternately, deletions in *H. cydno*. Nearly all of the intact genes displayed purifying selection (ω between 0.001 and 0.2459), with only the remaining alcohol dehydrogenase (ω = 1.2478) under positive selection (Table S5). Taken together, these results suggest that either of the alcohol dehydrogenase candidates may underlie the production of bioactive octadecanal from octadecanol in *H. melpomene*, although functional experiments are required to confirm this hypothesis. Alternately, as QTL for octadecanal and its likely precursor octadecanol overlap, a single FAR may be responsible for producing both volatiles.

Discussion

Previous work has shown that male *Heliconius* butterflies use pheromones during courtship, and the presence of these pheromones is necessary for successful mating to take place (Darragh et al 2017). Here we demonstrate that the two closely related species with strong mating differences, *H. melpomene* and *H. cydno*, show strong differences in chemical bouquets with a number of major compounds including octadecanal, 1-octadecanol, (*Z*)-11-icosenal, and (*Z*)-11-icosenol found in high concentrations in *H. melpomene* but largely absent in *H. cydno*, consistent with prior results (Mann et al 2017, Darragh et al 2019b). Somewhat surprisingly, despite these strongly divergent pheromone signals, there was little difference in antennal response to both overall wing chemical bouquets and various chemical

components between the species. Indeed, only octadecanal elicited a significant response in *H. melpomene*. Somewhat surprisingly, octadecanal was also physiologically active in *H. cydno*, where it is largely absent from the male wing bouquet.

These data stand in marked contrast to similar studies in other insect species and suggest that the peripheral nervous system architecture of *H. melpomene* and *H. cydno* has not diverged alongside their wing pheromone components. By contrast, the moth *Ostrinia nubialis*, whose *E*- and *Z*-strains diverged approximately 75,000-150,000 years ago (Malausa et al 2007), have opposite topologies in the antennal lobe and antennal sensillae (Kárpáti et al 2008, Koutroumpa et al 2014), and similar divergence in peripheral nervous system architecture has been seen in *Rhagoletis pomonella* (Frey & Bush 1990, Tait et al 2016) and *Drosophila mojavensis* (Date et al 2013, Crowley-Gall et al 2016) despite much shorter divergence times than between *H. melpomene* and *H. cydno*. Our results in *Heliconius* parallel what is seen in *Colias* butterflies, where *C. eurytheme* and *C. philodice* show very similar female electrophysiological responses to the con- and heterospecific pheromone compounds, despite a behavioral effect of treating males with heterospecific pheromones (Grula et al 1980). However, *Colias* use multiple signals when choosing between mates (Papke et al 2007), so rapid divergence in peripheral nervous system elements may not be necessary for full reproductive isolation. A similar pattern most likely occurs in *Heliconius*, with EAG responses very similar between species, and probable differences in the antennal lobe and higher brain regions (e.g. the mushroom body or lateral protocerebrum) driving behavioral differences (see e.g. Montgomery and Merrill 2017). Female *Heliconius* butterflies likely integrate multiple signals (including pheromones, male courtship flights, and visual cues) in these higher brain regions when making the decision to mate.

We have identified octadecanal as the major pheromone component in *H. melpomene* and showed that responses to it are conserved across both species despite its general absence in *H. cydno*. In the Lepidoptera, the activity of octadecanal as a pheromone has been tested in eight species across a variety of families (Tatsuki et al 1983; Cork et al 1988; Tumlinson et al 1989; Ho et al 1996; McElfresh et al 2000; Yildizhan et al 2009; El-Sayed et al 2011; Pires et al 2015; Chen et al 2018). Only in *Cerconota anonella*, and now in *Heliconius melpomene*, has some electrophysiological and behavioral activity been seen (Pires et al 2015). The closely related hexadecanal is a major pheromone component in the butterfly *Bicyclus anynana* (Nieberding 2008), and differs from octadecanal only in its origin from palmitic rather than stearic acid and carbon number, so octadecanal's role as a pheromone in *Heliconius* is perhaps not entirely unexpected from a biochemical perspective. In contrast, we have failed to

discover the physiologically active pheromone(s) in *H. cydno*. This is not entirely unexpected, as we focused on the most abundant compounds, but minor components have been shown to be important in other systems (McCormick et al 2014, Chen et al 2018).

Intriguingly, despite the strong EAG response, there was a marked negative behavioral response to increased octadecanal. A plausible explanation for the negative behavioral response to octadecanal supplementation is that disruption of the normal mixture ratios of *H. melpomene* may inhibit the female response, as seen in the butterfly *Pieris napi*, where synergistic processing of two volatile components in the male bouquet is necessary for acceptance behavior (Larsdotter-Mellström et al 2016). Octadecanal may also experience a dose-response curve with an aversive response to higher concentrations and an attractive response at lower ones. Potential mixture or dosage effects suggest that female *H. melpomene* may use octadecanal quantity or relative abundance to assess male quality or choose between courting males. The increased mating latency with octadecanal-treated males may reflect females undergoing a period of adjustment, either in the peripheral or central nervous system, to the higher dose of octadecanal; this would be consistent with our results showing long-term adaptation to octadecanal. We remain uncertain of what effect, if any, octadecanal would have on the behavior of *H. cydno*, where it is largely absent from the male pheromone bouquet. It may be used to avoid courtship with *H. melpomene*, supporting other divergent signals in maintaining reproductive isolation.

Given the strong physiological and behavioral response to octadecanal, and its possible role in reproductive isolation between *H. melpomene* and *H. cydno*, we sought to understand its genetic underpinnings. Fatty-acid derived compounds comprise the largest category of Lepidoptera sex pheromones (Ando & Yamakawa 2011), and are produced from fatty acyl-CoA precursors via the action of several enzymes. Since these pheromones are secondary metabolites derived from primary metabolic pathways, their production is likely to be relatively labile in evolutionary terms, allowing simple genetic changes to drive the wide diversity of lepidopteran sex pheromones. Even though we have a broad knowledge of pheromone diversity in Lepidoptera, our understanding of the genetics of pheromone biosynthesis is relatively weak. To date, only a handful of genes have been identified, mostly desaturases and fatty acyl-CoA reductases (Yew & Chung 2015).

Using a QTL mapping approach, we have shown that the production of octadecanal has a relatively simple genetic basis, with a region on chromosome 20 corresponding to production of both octadecanal and its likely precursor octadecanol in *Heliconius*. This locus therefore likely represents a region under divergent selection between *H. melpomene* and *H. cydno* that is unlinked to previously identified species

differences in color and mate choice (Jiggins 2017, Merrill et al 2019). Patterns of F_{ST} between the species are highly heterogeneous and were not especially informative in further delimiting the locus (data from Martin et al 2013). Due to our small mapping population, the confidence intervals for these QTL therefore remain large: the octadecanal QTL spans 3.4Mb and contains 160 genes. Of these, we identified 16 likely candidate genes based on known biosynthetic pathways in moths and the butterfly *Bicyclus anynana* (Lienard et al 2014): 14 fatty acyl-CoA reductases and two alcohol dehydrogenases. Fatty acyl-CoA reductases have previously been identified in *H. melpomene* by (Lienard et al 2014), who noted lineage-specific duplications within *H. melpomene* on two scaffolds corresponding to *H. melpomene* chromosomes 19 and 20. All but one of the candidate FARs found on chromosome 20 were identified by Lienard et al, but all fall outside their clade of pheromone gland FARs. The identified *Bicyclus* FAR that produces hexadecanol does not also produce the major pheromone hexadecanal, implying the presence of an additional as yet undescribed alcohol dehydrogenase in *Bicyclus*, and the biochemical similarity between hexadecanal and octadecanal suggests *Heliconius* may also use an alcohol dehydrogenase to produce octadecanal. By contrast, the overlapping octadecanol and octadecanal QTL on chromosome 20 in *Heliconius* suggest the presence of a bifunctional FAR that produces both the alcohol and aldehyde together. Further studies, including functional assays and location of wing pheromone biosynthesis, will be required to tease apart our potential candidates.

Our studies of the electrophysiological and behavioral responses of *Heliconius* butterflies and the genetic basis of pheromone production add to the growing body of literature suggesting that pheromonal communication in Lepidoptera is not limited to nocturnal moths but can be found in day-flying butterflies that also use striking visual signals. *Heliconius* butterflies can detect con- and heterospecific wing compound bouquets, and a major component, octadecanal, is physiologically and behaviorally active in *H. melpomene* and its genetic basis appears relatively simple, consistent with other pheromone shifts found in insects (Symonds & Elgar 2007). Along with their striking wing color patterns, male *Heliconius* use chemistry to affect female mate choice, combining courtship behaviors, and chemistry in a dance to elicit female mating responses (Klein & de Araújo 2010, Mérot et al 2015). Despite our human bias towards visual signals, we are now beginning to understand how such visually striking butterflies communicate using chemistry.

Methods

Data and analysis code are deposited in Dryad under accession XXX. Sequencing reads leading to linkage map construction are deposited in the European Nucleotide Archive (ENA) under accession YYY.

Butterflies

Stocks from central Panama of *Heliconius melpomene rosina* and *H. cydno chioneus* (hereafter *H. melpomene* and *H. cydno*) were used for all experiments (Darragh et al 2017). Butterflies were reared in insectaries at the Smithsonian Tropical Research Institute, Gamboa, Panama under ambient temperature and light conditions. Eggs were collected from these breeding stocks and the resulting larvae fed on *Passiflora platyloba var. williamsi*, *P. biflora*, *P. menispermifolia*, and *P. vitifolia* until pupation. Data from (Darragh et al 2019a) indicate that larval diet does not affect the major compounds found in *H. melpomene*, suggesting that this dietary variation is unlikely to affect results. Newly-eclosed adult butterflies were separated by sex to ensure virgin status and supplied with flowers from *Psiguria warscewiczii*, *Psiguria triphylla*, *Gurania eriantha*, *Psychotria poeppigiana*, *Stachytarpheta mutabilis*, and *Lantana* sp. (most likely *L. camara*) as pollen sources, as well as a ~20% sucrose solution as a nectar source. All experiments used virgin butterflies. For assessment of *H. cydno* wing bouquets, male butterflies were between 10-12 days post eclosion and had fed exclusively on *Passiflora platyloba var. williamsi*. For electrophysiology, female butterflies were between 1 and 20 days post eclosion, and males between 10-20 days post eclosion to ensure male sexual maturity (Darragh et al 2017). For behavior, female butterflies were used the day after eclosion and males were between 10 and 59 days post eclosion. Natural wing extracts of both species were extracted from males 10-12 days post eclosion as described in Darragh et al. 2017 using dichloromethane plus 1 ng/µL 2-tetradecyl acetate (hereafter "DCM+IS") and concentrated approximately 10x prior to use under room air. All samples were stored at -20°C before use.

Identification and quantification of androconial compounds

To identify species specific compounds among our two species, the chemical composition of the *H. cydno* androconial bouquet was investigated in samples from 26 adult male *H. cydno* and compared with 31 adult male *H. melpomene*, the latter including samples previously analyzed in (Darragh et al 2017, 2019a), all collected as above. Samples were assessed using gas chromatography-mass spectrometry (GC-MS) with an Agilent 7890B gas chromatograph coupled with an Agilent 5977 mass spectrometer with electron ionization (Agilent Technologies, California, USA). The GC utilized an Agilent HP-5MS capillary column (30m length, 0.25mm inner diameter), helium carrier gas at 1.2 mL/minute, and an

Agilent ALS 7693 autosampler. Injection was splitless with an inlet temperature of 250°C. The temperature ramp was isothermal at 50°C for 5 minutes, then increased at 5°C/minute to 320°C and was then isothermal for 5 minutes. Samples were identified using a custom MS library and quantified by comparison with the internal standard.

In line with (Darragh et al 2017), wings from eight male *H. cydno* were dissected into four regions: hindwing androconia, forewing overlap region, hindwing rest-of-wing, and forewing rest-of-wing, and all extracted identically after dissection. Wing region area was quantified by photographing the wings before dissection and measuring the total pixel area of each wing region in the GNU Image Manipulation Program v.2.8.20 (GIMP Development Team), with the pixel-mm conversion via measurement of a ruler in the photograph. Quantified compounds in each wing region for each individual were scaled by the area of the relevant wing region in that individual.

Chemicals

Syringaldehyde (3,5-dimethoxy-4-hydroxybenzaldehyde), 1-octadecanol, and henicosane were obtained commercially (Sigma-Aldrich). The aldehydes octadecanal, (*Z*)-11-icosenal, and (*Z*)-13-docosenal were obtained from the respective alcohols 1-octadecanol, (*Z*)-11-icosen-1-ol, and (*Z*)-13-docosen-1-ol by oxidation with iodoxybenzoic acid (IBX) in ethyl acetate according to More and Finney (2002). The required alcohols (*Z*)-11-icosen-1-ol and (*Z*)-13-docosen-1-ol were in turn obtained by lithium aluminum hydride reduction from commercially available (*Z*)-11-icosenoic acid and methyl (*Z*)-13-docosenoate (Larodan) according to Cha and Brown (1993). The seven target compounds (see SI for structures and reaction scheme) were chosen due to their quantitative dominance in the chemical profiles of *H. melpomene* and *H. cydno*. The solvent for all synthesized compounds was hexane, with the exception of the polar syringaldehyde, which was diluted in a 1:10 mixture of dichloromethane and hexane.

Synthetic blends of *H. melpomene* and *H. cydno* male wing bouquets were prepared from these synthesized compounds. Synthetic *H. melpomene* contained 23.22 ng/µL syringaldehyde, 23.33 ng/µL octadecanal, 6.92 ng/µL 1-octadecanol, 4.67 ng/µL (*Z*)-11-icosenal, 20.33 ng/µL (*Z*)-11-icosenol, and 4.83 ng/µL (*Z*)-13-docosenal. Synthetic *H. cydno* contained 47 ng/µL syringaldehyde and 93.33 ng/µL henicosane. Floral direct extractions from *Lantana* sp. (most likely *L. camara*) growing wild in Gamboa, Panama were used as a positive control. Single umbels were removed from plants at dawn and placed in a scintillation vial to which 400uL of DCM+IS was added. After 1 hour, the DCM+IS was removed to a glass vial and kept at -20°C before use.

355 Electroantennography

356 Electrophysiological preparations were assembled as follows: antennae were cut from the head of a
357 virgin butterfly using fine scissors and the final 6.5 segments (approximately 1.5mm) cut off with a
358 scalpel. Both antennae were then placed in parallel across an antenna fork (Syntech, Buchenbach,
359 Germany) and held in place using electrode gel (Spectra 360 Electrode Gel, Parker Laboratories Inc.,
360 Fairfield, NJ, USA). The antenna fork was mounted on a Syntech EAG CombiProbe with 10x internal gain,
361 and signal from this was routed through a Syntech IDAC4. EAG waveforms recorded using Syntech
362 GcEad/2014 software. Stimulus pulses were delivered using a Syntech CS-55 Stimulus Controller with
363 foot pedal trigger. Both continuous unhumified room air and stimulus pulses were delivered to the
364 preparation at 1.5 liters/min. The stimulus pulses were delivered in triplets of 0.5 seconds each,
365 separated by 5 seconds, with triplets initiated every 30 seconds. Stimulus delivery used odor cartridges
366 assembled from Pasteur pipettes with a strip of filter paper plugged with cotton when not in use; each
367 stimulus cartridge was 'charged' with 10uL of stimulus solution for each experiment. Each antennal
368 preparation was used only once.

369 Two sets of stimuli were delivered: a species comparison set and a synthetic compound set. Both sets
370 used air (nothing added to filter paper), hexane, and DCM+IS as negative controls and *Lantana* extract as
371 a positive control. The species comparison set included male wing extracts from *H. melpomene* and *H.*
372 *cydno* ("Mnat" and "Cnat" respectively) and synthetic blends representing the two species ("Msyn" and
373 Csyn"). The synthetic compound set included air, hexane, DCM+IS, the conspecific male wing extract,
374 the conspecific synthetic blend, and the seven synthetic compounds. Presentation order was
375 randomized before each experiment. Species comparison experiments consisted of sixteen pulses each
376 of the seven stimuli, interspersed with five pulses of *Lantana* extract at the start, between every three
377 stimuli, and at the end. Synthetic compounds were similar, with eleven pulses of each of the twelve
378 stimuli, interspersed with four pulses of *Lantana* at the start, between every four stimuli, and at the end.
379 For analysis, the first triplet of each stimulus set was removed, leaving fifteen and ten pulses
380 respectively. At least ten female and five male butterflies of each species were used with each
381 experiment.

382 Onset time and amplitude of EAG responses were marked using GcEad/2014. To control for antennal
383 signal degradation over time, a time-dependent correction factor was calculated using linear
384 interpolation between each *Lantana* set and this applied to the absolute value of the EAG response

amplitude. These corrected amplitudes were then scaled to the amplitude of the initial *Lantana* set to partially control for differences between preparations.

Analysis of antennal adaptation

Short-term adaptation (STA), as defined in (Zufall & Leinders-Zufall 2000), is adaptation occurring only over a very brief window from the initial stimulus that then resolves quickly with no further stimulation (e.g. the 10 second interval given for salamanders). By contrast, long-term or long-lasting adaptation (LTA) are defined in the same source as persisting over an extended period of time, up to several minutes in vertebrates and insects (Stengl 2010). Within our electrophysiological data set, there is the potential to measure both types of adaptation. STA, if present, should be evident within an individual triplet, as the stimuli within a triplet are separated by 5 seconds (and thus STA is likely to persist within a triplet), whereas LTA should be evident across an individual stimulus set, as these lasted 5.5-8 minutes with maximum intervals of 30 seconds between stimuli, insufficient for LTA to be abolished if we assume similar mechanisms as vertebrates. We assessed antennal responses for LTA by pooling all triplets within a stimulus-species-sex combination. We pooled these data within stimulus set types (species comparison and synthetic compound), treating butterfly preparation identity as a random effect. For each stimulus, the change in antennal response was assessed over the time since initial presentation of the stimulus. STA was assessed by looking for significant changes in the residuals from this analysis between members of a triplet.

Behavior

To test the potential role of octadecanal in *H. melpomene* female mate choice, behavioral experiments were conducted in insectaries at STRI, Gamboa, Panama, between April and July 2018. One day old virgin females were presented with both an octadecanal treated and a control *H. melpomene* male for two hours. Males were at least ten days old and were selected based on similarity of size, wing-wear and age, with treatment being allocated randomly by coin flip. Either 25 µl octadecanal (140 ng/µl octadecanal in hexane, thus adding 3500ng to the existing average 773.4ng for approximately a 5.5x dose) or 25 µl pure hexane (both evaporated down to a smaller volume under room air) was applied to the hindwing androconial region of each male, and males were then allowed 30 minutes to settle before beginning the two hour experiment period. Experiments began at or close to 9am, with observations being made every 15 minutes or until mating occurred.

To test the persistence of the octadecanal treatment on the wings of live butterflies, a separate set of *H. melpomene* males was treated as above with either hexane or octadecanal. Separate males were sampled at 30 minutes post treatment and two hours post treatment by extraction of the forewing overlap region (Darragh 2017) and the hindwing androconia in DCM+IS as above, with two males per treatment-time combination. Octadecanal was then measured using GC-MS as above and quantified by peak area comparison with the 2-tetradecyl acetate internal standard.

Quantitative trait locus mapping for octadecanal production

To map the genetic basis for octadecanal production in *H. melpomene*, we took advantage of the fact that *H. cydno* produces little to no octadecanal. Bidirectional F_1 crosses between the two species revealed that the *H. cydno* phenotype (low to no octadecanal) is dominant over the high octadecanal production found in *H. melpomene*, so we constructed backcross families by crossing F_1 males to female *H. melpomene* from our existing stocks. A total of ten families (nine with a female *H. melpomene* grandparent and one with a female *H. cydno* grandparent) were constructed, with each offspring representing a single recombination event from the F_1 father. Butterflies were reared and wing extracts collected and analyzed from male offspring as described above, except that all larvae were reared on *Passiflora platyloba var. williamsi*. Bodies of male offspring were collected into dimethyl sulfoxide (DMSO) for later library preparation. The Castle-Wright estimators for octadecanal and octadecanol production were calculated using the phenotypic variance of the backcross individuals as the estimated segregation variance (Jones 2001).

Qiagen DNeasy kits (Qiagen) were used for DNA extraction. Individuals were genotyped either by RAD-sequencing as previously described (Davey et al., 2017, Merrill et al., 2019), or low-coverage whole genome sequencing using nextera-based libraries. For the nextera-based libraries a secondary purification was performed using magnetic SpeedBeads™ (Sigma) dissolved in 0.44mM PEG8000, 2.5M NaCl, 1mM Tris-Cl pH=8, and 0.1mM EDTA pH=8.0. High-throughput sequencing libraries were generated using a method based on the Nextera DNA Library Prep (Illumina, Inc.) with purified Tn5 transposase (Picelli et al 2014). Sample barcoding was performed using PCR extension with an i7-index primer (N701–N783) and the N501 i5-index primer. Libraries were purified and size selected using the same beads as above. Pooled libraries were sequenced by HiSeq 3000 (Illumina) by BGI (China).

Linkage mapping was conducted using standard Lep-MAP3(LM3) pipeline (Rastas, 2017) . First, individual fastq files were mapped to the melpomene reference genome using BWA MEM (Li, 2013) and

then sorted bams were created using SAMtools (Li & Durbin, 2011). The input genotype likelihoods were constructed by SAMtools mpileup and pileupParser2+pileup2posterior from LM3. The pedigree of individuals was validated and corrected using IBD (identity-by-descent) module and the sex of individuals was validated and corrected according to the coverage on the Z chromosome and autosomes using SAMtools depth. Then, ParentCall2 (parameter ZLimit=2) and Filtering2 (dataTolerance=0.001) modules were called on the input data and a random subset of 25% of markers (to speed up analysis) was used for the subsequent steps.

Initial linkage groups (chrX.map) and marker orders (orderX.txt) were constructed based on melpomene genome for each of 21 chromosomes. SeparateChromosomes2 was run on each of these groups with the default parameters (lodLimit=10) except for map=chrX.map (for X=1..21). Finally, OrderMarkers2 was run on each chromosome in the constructed order, with parameter scale=0.05, recombination2=0, evaluateOrder=orderX.txt and map=result_from_SeparateChromsomes2.chrX.txt. Another evaluation was done with data in the grandparental phase (additional parameter grandparentPhase=1). The phased data of these orders were matched using phasematch script (LM3) and obtaining all markers from the first evaluation in the grandparental phase. This obtained result was used as the final map.

Map construction resulted in the retention of 447,818 SNP markers across 89 individuals with phenotype data. To facilitate computation, markers were thinned evenly by a factor of ten, resulting in 44,782 markers with no missing data. Octadecanal production was log-transformed to obtain normality, then regressed against marker position using the R/qtl2 R library (Broman et al 2018). Significance thresholds were obtained by permutation testing in R/qtl2 with 1000 permutations, and QTL confidence intervals obtained using the bayes_int command. To account for the family structure present in our QTL mapping populations, we additionally included a kinship matrix calculated by R/qtl2 using the LOCO (leave one chromosome out) method in the marker regression and recalculated significance thresholds and confidence intervals with the kinship term included. Percent variance explained was calculated as the difference in phenotype means of individuals of each genotype divided by the difference in the parental phenotype. Since the genetic linkage map was based on whole genome data, we were able to obtain physical positions of QTL confidence interval endpoints. The physical positions of the kinship-included confidence interval were used to query Lepbase (Challis et al 2016) for potential candidate genes from the *H. melpomene* genome. To identify putative functions for each potential candidate, protein sequences from Lepbase were searched against the nr (non-redundant) protein database using BLASTp (Altschul et al 1990). For each candidate with a promising functional annotation, exons were

pulled out of the *H. cydno* genome (Pessoa Pinharanda 2017) and aligned to the *H. melpomene* genes using BLASTn with each *H. melpomene* exon to search for SNPs between the two species. Selection was tested at the sequence level using codeml (Yang 2007) in pairwise mode with the F3X4 codon frequency model.

Statistical analysis

Differences in individual compounds between *H. melpomene* and *H. cydno* were assessed using a Welch's *t*-test. Wing region differences in *H. cydno* were assessed for each individual compound found in at least four of eight samples of at least one wing region using a linear mixed model with wing region as a fixed effect and butterfly identity as a random effect using the package nlme v.3.1.137 (Pinheiro et al 2018). Statistical tests for these models were assessed with the Anova function in the package car v.3.0.0 (Fox & Weisberg 2011), and comparisons between wing regions were performed using the emmeans package v.1.3.1 (Lenth 2018).

For electroantennography, species comparison sets and synthetic compound sets were analyzed separately. Corrected and scaled EAG responses (for each experiment within sexes and species) were compared between stimuli with a linear mixed model for *H. melpomene* with stimulus as a fixed effect and butterfly preparation identity as a random effect using nlme as above. EAG responses for *H. cydno* were better represented as an inverse gamma distribution, so were assessed in the same fashion with a generalized linear mixed model (glmmPQL) using the package MASS v.7.3.50 (Venables & Ripley 2002). Statistical tests for these models were assessed with the Anova and emmeans (version 1.1.3) functions as above.

Long-term adaptation was assessed using robust linear mixed models with the package robustlmm v.2.2.1 (Koller 2016), with corrected but unscaled amplitude as the response variable, time since initial presentation of the stimulus as a fixed variable, and butterfly preparation as a random variable. Responses were pooled across samples within a species-sex combination and considered for each stimulus separately. As robustlmm does not provide p-values, confidence intervals were used to assess significance and difference between sample-species combinations. Short-term adaptation was assessed using the residuals from the same regression. Differences in the residuals between triplets within a triplet set were tested using a one-sample t-test with a hypothesized mean value of zero (i.e. no difference between residuals), performed on the subtractive difference between the residuals of the third (last) and first triplets.

505 Female mate choice was assessed using a binomial test, and treatment differences in time until mating
506 were assessed with a two-sided t-test assuming unequal variances. Octadecanal persistence was not
507 assessed statistically due to the small sample size. All statistical tests were performed in R v.3.5.0,
508 v.3.5.1, or v.3.5.2 (R Core Team 2013).

509

510 **Acknowledgments**

511 The authors wish to thank Bill Wcislo and Callum Kingwell for advice on electroantennography. We also
512 thank Ian Warren, the insectary team in Panama including Oscar Paneso and Chi-Yun Kuo, as well as the
513 administrative support of Oris Acevedo and Melissa Cano. We also thank S. Dilek for technical
514 assistance. Permits for research and collection of butterfly stocks were provided by the government of
515 Panama. KJRPB, KD, JM, IAW, and CDJ were funded by the European Research Council (FP7-IDEAS-ERC
516 339873); KD was additionally funded by a Natural Research Council Doctoral Training Partnership and a
517 Smithsonian Tropical Research Institute Short Term Fellowship; YFC and MK are supported by the
518 European Research Council Grant No. 639096 "HybridMiX" and by the Max Planck Society; WOM is
519 funded by the Smithsonian Tropical Research Institute; and SS by the Deutsche Forschungsgemeinschaft
520 (grant DFG Schu984/13-1).

521 **Author contributions**

522 KJRPB designed, performed, and analyzed electrophysiology experiments, performed QTL mapping and
523 downstream analysis, and wrote the manuscript. KD designed and performed QTL mapping crosses,
524 designed behavioral experiments, and prepared sequencing libraries. KD, SFG, DAA, KJRPB, and JM
525 reared butterflies for all experiments. JM performed and analyzed behavioral experiments. IAW
526 prepared sequencing libraries. YFC and MK provided Tn5 transposase for sequencing. SS provided
527 synthetic compounds used in electrophysiology and behavior experiments as well as mass spectrometry
528 libraries. PR analyzed sequencing data and constructed linkage maps. RMM helped design the QTL
529 experiment. CDJ and WOM contributed funding, resources, and helped design the overall study. All
530 authors contributed to the manuscript editing process.

531 **Tables**

532 Table 1: Electroantennographic responses of the different sex-species combinations to individual
533 compounds compared with hexane (P2-7) or hexane and dichloromethane (P1). Only significant

alterations are shown. No asterisk: $p \leq 0.05$; *: $p \leq 0.01$; **: $p \leq 0.001$; ***: $p \leq 0.0001$. P1: syringaldehyde, P2: octadecanal, P3: 1-octadecanol, P4: henicosane, P5: (*Z*)-11-icosenal, P6: (*Z*)-11-icosenol, P7: (*Z*)-13-docosenal. *Heliconius melpomene* wing bouquet contains P1-7; *H. cydno* wing bouquet contains P1, P4, and P6 (Table S1). Artificial *H. melpomene* wing blend contains the dominant compounds P1-P3 and P5-P7; artificial *H. cydno* wing blend contains dominant compounds P1 and P4.

Species	**Sex**	**n**	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**
H. melpomene	Female	11	-	↑***	-	-	-	-	-
H. melpomene	Male	5	-	↑***	-	-	↓**	-	-
H. cydno	Female	10	-	↑***	-	-	↓	↓***	↓
H. cydno	Male	5	-	↑***	↑*	↑***	-	↑*	-

SI Table 1: Comparison of *Heliconius cydno* and *H. melpomene* wing androconia bouquet. Amounts given are the mean in nanograms across 31 (*H. melpomene*) and 26 (*H. cydno*) samples, ± SD. Compounds are only included if found in at least 1/3 of samples for at least one species at levels of at least 1ng. #: compound found in at least 90% of samples of that species. Bold: species differ significantly. dof: Welch's two sample t-test degrees of freedom.

SI Table 2: Wing region specificity of *Heliconius cydno* androconial compounds. Amounts given are the mean across eight samples ± SD. Numbers in parentheses after compound amounts indicate the number of samples the compound was found in. Bold: wing regions differ significantly. NA: post-hoc test not conducted as original linear model not significant.

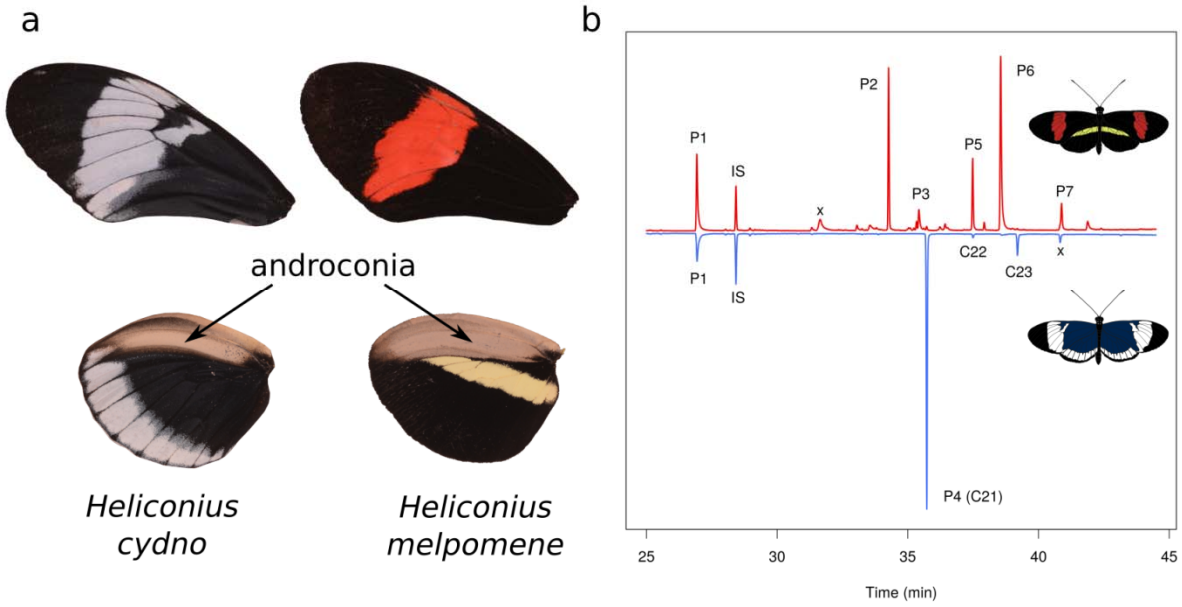
SI Table 3: Details of statistical comparisons of stimuli from electrophysiological experiments. P1-P7: as in the main text and Figure 4.

SI Table 4: Short-term adaptation to different stimuli in *H. melpomene* and *H. cydno* males and females. t1, t2, and t3: first, second, and third members of a stimulus triplet. dof: degrees of freedom. P1-P7: as in the main text and Figure 4.

SI Table 5: Potential candidate genes for octadecanal production underlying the QTL peak on chromosome 20 in *Heliconius melpomene*.

Figures

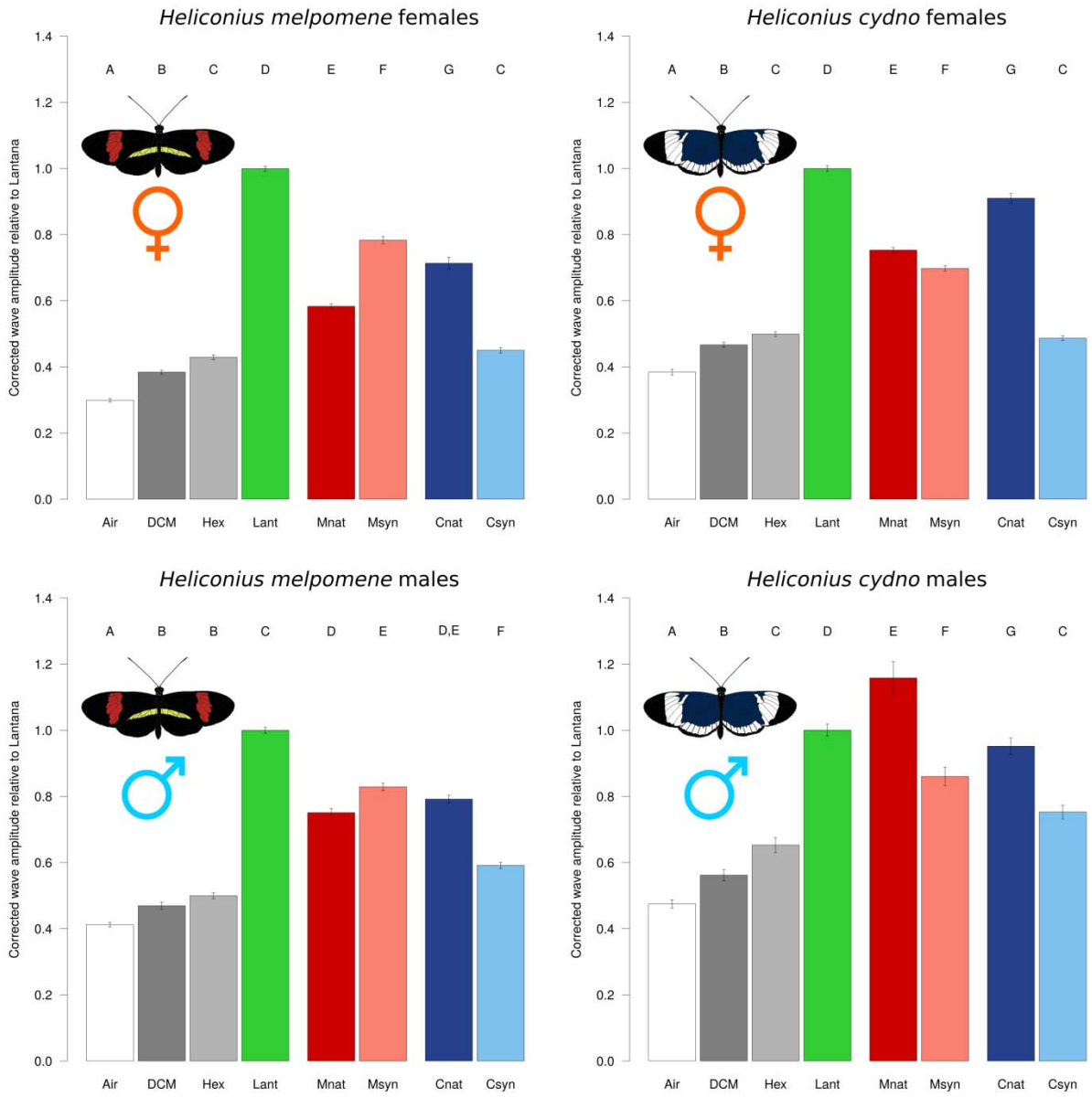
557 Figure 1: Chemical ecology of *Heliconius melpomene* and *H. cydno*. a: dorsal forewing and hindwing of
558 each species showing the silvery androconial region of the hindwing used during male courtship. b: Total
559 ion chromatogram of *H. melpomene* (top) and *H. cydno* (bottom) wing androconia. P1: syringaldehyde;
560 P2: octadecanal; P3: 1-octadecanol; P4: henicosane; P5: (*Z*)-11-icosenal; P6: (*Z*)-11-icosenol; P7: (*Z*)-13-
561 docosenal; IS: internal standard (2-tetradecyl acetate); x: contaminant; C21: henicosane; C22: docosane;
562 C23: tricosane.



563

564

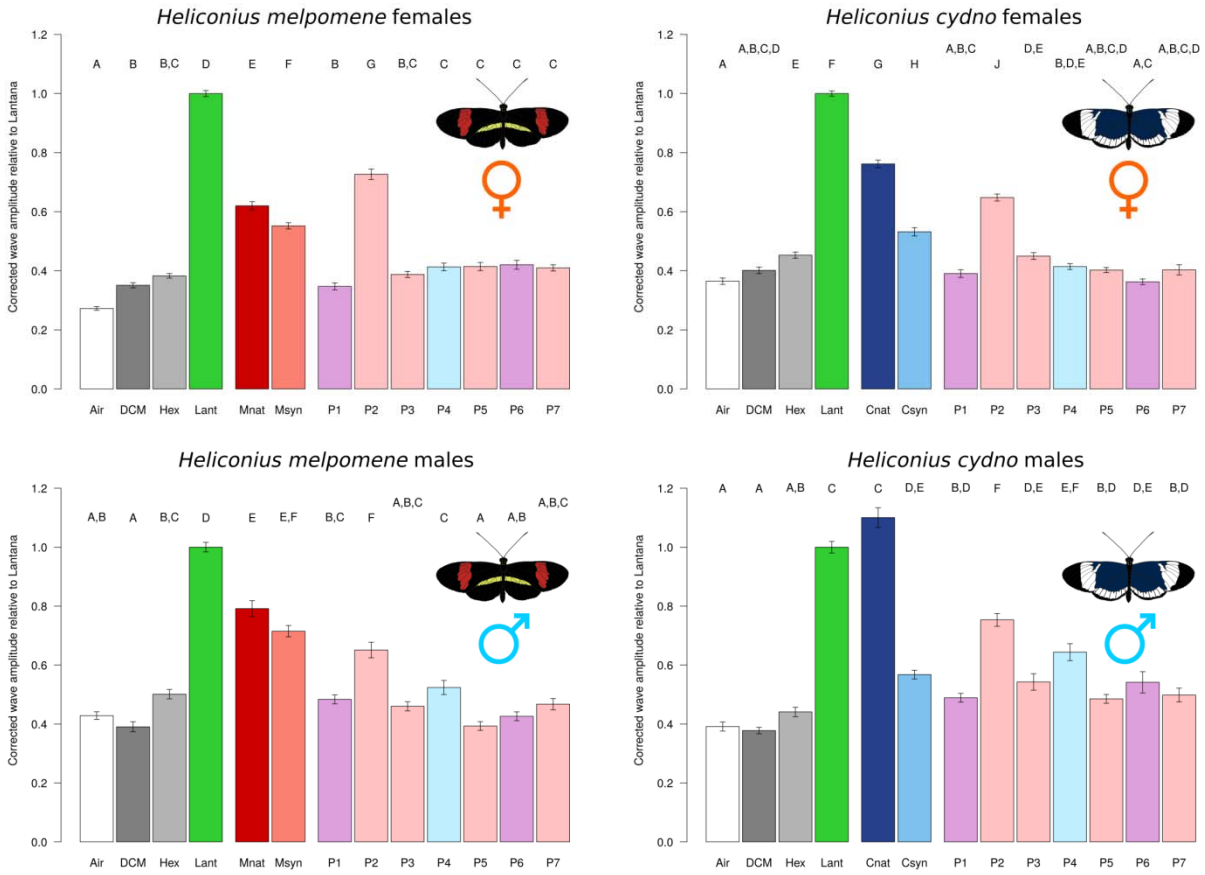
565 Figure 2: Electroantennographic responses of *Heliconius* butterflies to conspecific and heterospecific
566 wing extracts. Stimuli: air, dichloromethane plus 2-tetradecyl acetate (internal standard), hexane,
567 *Lantana* extract, natural male *H. melpomene* wing extract, synthetic *H. melpomene* blend, natural male
568 *H. cydno* wing extract, synthetic *H. cydno* blend. Bars: average of normalized corrected amplitude ± SEM.



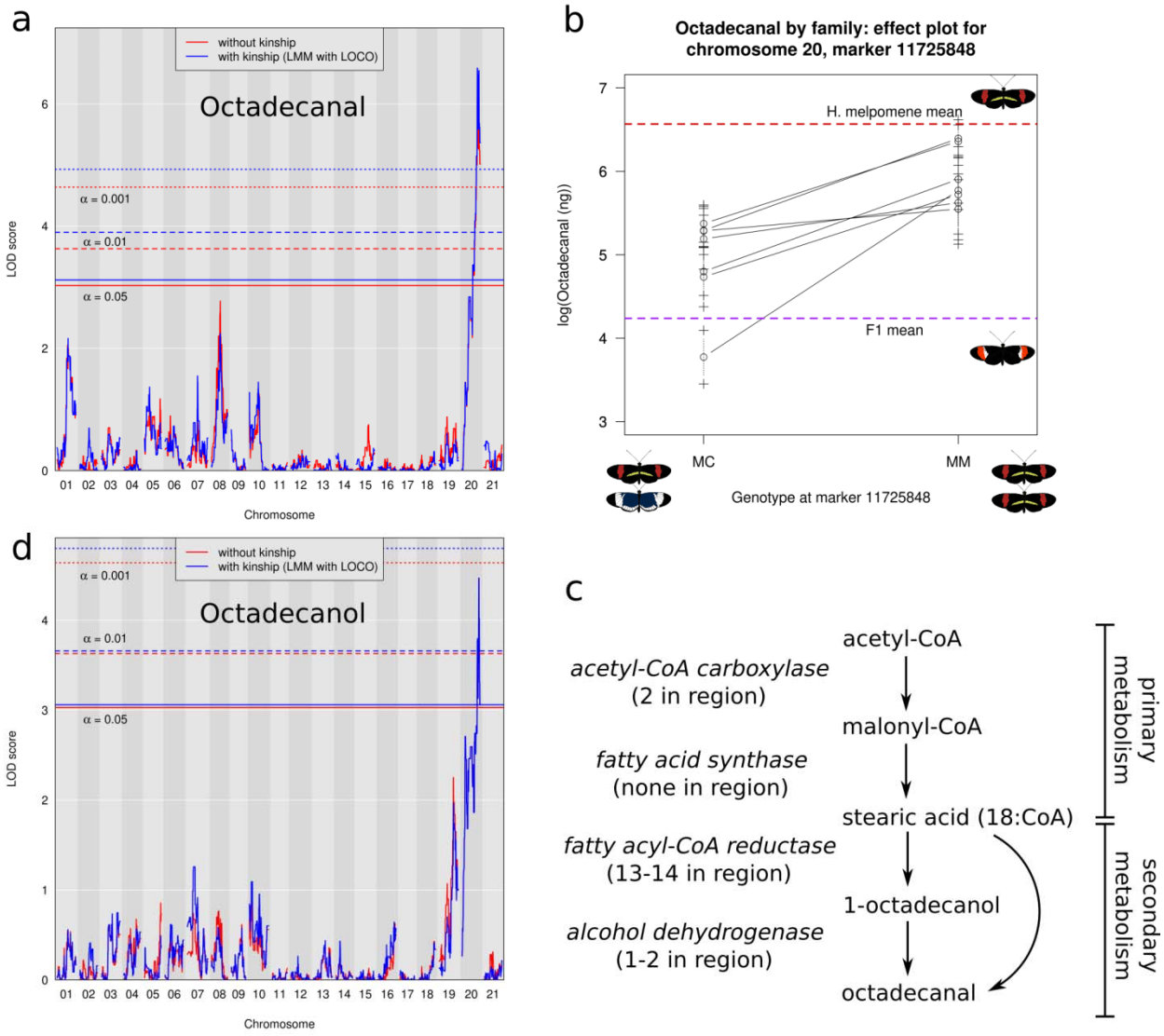
569

570

Figure 3: Electroantennographic responses of *Heliconius* butterflies to synthetic wing compounds. Stimuli: air, dichloromethane plus 2-tetradecyl acetate (internal standard), hexane, *Lantana* extract, natural male *H. melpomene* or *H. cydno* wing extract, synthetic *H. melpomene* or *H. cydno* blend. P1-P7 as in Figure 1. Bars: average of normalized corrected amplitude ± SEM. Light red/light blue: compound found in all *H. melpomene* (red) or *H. cydno* (blue) and >1% of natural bouquet makeup; purple: compound found in both species and >1% of natural bouquet makeup in at least one species.



579 Figure 4: QTL mapping of octadecanal and octadecanol production in *Heliconius melpomene*. a: QTL map
580 for production of octadecanal. b: Effect plots at the peak of the locus on chromosome 20 for the seven
581 individual backcross mapping families with at least 5 individuals. c: Potential biosynthetic pathway for
582 octadecanal production. d: QTL map for production of octadecanol, a potential precursor of
583 octadecanal.



584

585

586 SI Figure 1: Comparison of the major compounds in *Heliconius melpomene* and *H. cydno*. Compounds
587 shown are those contributing at least 1% of the total bouquet amount in either species. Numbers under
588 each bar indicate how many samples (out of 31 for *H. melpomene* and 26 for *H. cydno*) the compound
589 was found in. Significantly different compounds: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. n.s., not
590 significant.

591 SI Figure 2: Absolute and relative abundance of different compound classes in *H. melpomene* and *H.*
592 *cydno*. n.s., not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. The two unknown categories were not
593 tested as the compound types are not known.

594 SI Figure 3: The seven compounds found in at least 0.1ng/mm^2 of wing tissue in at least one wing region
595 in *Heliconius cydno*. a: Compound abundance per square millimeter of tissue. b: Compound abundance
596 without tissue area correction. Numbers under each bar indicate how many samples (out of eight) the
597 compound was found in; letters above bars indicate significant differences between regions. n.s., not
598 significant. A: hindwing androconia; O: forewing overlap region; H: hindwing excluding androconia; F:
599 forewing excluding overlap region.

600 SI Figure 4: Structures and names of major components of the androconia of *H. melpomene* and *H.*
601 *cydno* used in electrophysiological experiments.

602 SI Figure 5: Synthesis of target compounds used in electrophysiological experiments. IBX: iodosobenzoic
603 acid; LiAlH: lithium aluminum hydride.

604 SI Figure 6: *Heliconius melpomene* responds to electrophysiological stimuli. Top to bottom:
605 dichloromethane plus 2-tetradecyl acetate (internal standard) (negative control), *Lantana* extract
606 (positive control), natural *H. melpomene* male wing extract, depiction of stimulus pulse timing. Data
607 from a single virgin female.

608 SI Figure 7: Long-term adaptation to natural and synthetic stimuli in *Heliconius* butterflies. The 95%
609 confidence intervals of the robust LMM slope are shown; a negative slope means that responses to that
610 stimulus drop over time. P1-P7: see Figure 2.

611 SI Figure 8: Strength of long-term adaptation correlates with amplitude of EAG response in a sex-specific
612 fashion. In females, a stronger response to a given stimulus correlates with a stronger degree of LTA
613 both overall and for the synthetic compound set. In males the same is seen overall and for the natural
614 extract set in *H. cydno*, but not in *H. melpomene*.

615 SI Figure 9: Octadecanal persistence in treated males over time. Bars show individual males, with two
616 males per treatment-time point combination. Hindwing androconia and forewing overlap region males
617 are shown in the same positions.

618 SI Figure 10: Octadecanal in *H. melpomene*, *H. cydno*, two F_1 families (one in each crossing direction),
619 and the ten backcross to *H. melpomene* families used in QTL mapping.

620 SI Figure 11: Chromosome 20 QTL map for production of octadecanal and octadecanol in *H. melpomene*.
621 Shaded regions indicate the Bayesian confidence intervals with kinship structure taken into account and
622 black line indicates the peak of the QTL.

623

624 **Literature cited**

625 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol*
626 **215**:403-410.
627
628 Andersson J, Borg-Karlson A-K, Vongvanich N, Wiklund C (2007) Male sex pheromone release and female
629 mate choice in a butterfly. *J Exp Biol* **210**:964-970.
630
631 Ando T, Yamakawa R (2011) Analyses of lepidopteran sex pheromones by mass spectrometry. *Trends*
632 *Anal Chem* **30**(7):990-1002.
633
634 Broman KW, Gatti DM, Simecek P, Furlotte NA, Prins P, Sen Ś, Yandell BS, Churchill GA (2018) R/qtl2:
635 software for mapping quantitative trait loci with high-dimensional data and multi-parent populations.
636 *Genetics* **211**:495-502.
637
638 Cha, J. S., and Brown, H. C. 1993 Reaction of sodium aluminum hydride with selected organic
639 compounds containing representative functional groups. Comparison of the reducing characteristics of
640 lithium and sodium aluminum hydrides. *J Org Chem* **58**:4727–4731.

641 Challis RJ, Kumar S, Dasmahapatra KK, Jiggins CD, Blaxter M (2016) Lepbase: the Lepidopteran genome
642 database. *bioRxiv* 056994.
643
644 Chen Q-H, Zhu F, Tian Z, Zhang W-M, Guo R, Liu W, Pan L, Du Y (2018) Minor components play an
645 important role in interspecific recognition of insects: a basis to pheromone based electronic monitoring
646 tools for rice pests. *Insects* **9**(4):192.
647
648 Cork A, Chamberlain DJ, Beevor PS, Hall DR, Nesbitt BF, Campion DG, Attique MR (1988) Components of
649 female sex pheromone of spotted bollworm, *Earias vittella* F. (Lepidoptera: Noctuidae): identification
650 and field evaluation in Pakistan. *J Chem Ecol* **14**(3):929-945.
651

652 Crowley-Gall A, Date P, Han C, Rhodes N, Andolfatto P, Layne JE, Rollmann SM (2016) Population
653 differences in olfaction accompany host shift in *Drosophila mojavensis*. *Proc R Soc B* **283**:20161562.
654
655 Darragh K, Byers KJRP, Merrill RM, McMillan WO, Schulz S, Jiggins CD (2019a) Male pheromone
656 composition depends on larval but not adult diet in *Heliconius melpomene*. *Ecol Entomol* **44**(3):397-405.
657
658 Darragh K, Montejo-Kovacevich G, Kozak KM, Morrison CR, Figueiredo CME, Ready JS, Salazar C, Linares
659 M, Byers KJRP, Merrill RM, McMillan WO, Schulz S, Jiggins CD (2019b) Species specificity and
660 intraspecific variation in the chemical profiles of Heliconius butterflies across a large geographic range.
661 *bioRxiv* 573469.
662
663 Darragh K, Vanjari S, Mann F, Gonzalez-Rojas MF, Morrison CR, Salazar C, Pardo-Diaz C, Merrill RM,
664 McMillan WO, Schulz S, Jiggins CD (2017) Male sex pheromone components in *Heliconius* butterflies
665 released by the androconia affect female choice. *PeerJ* **5**:e3953.
666
667 Date P, Dweck HKM, Stensmyr MC, Shann J, Hansson BS, Rollmann SM (2013) Divergence in olfactory
668 host plant preference in *D. mojavensis* in response to cactus host use. *PLOS ONE* **8**(7):e70027.
669
670 John W Davey, Sarah L Barker, Pasi M Rastas, Ana Pinharanda, Simon H Martin, Richard Durbin, W Owen
671 McMillan, Richard M Merrill, Chris D Jiggins (2017) No evidence for maintenance of a sympatric
672 Heliconius species barrier by chromosomal inversions. Evol Lett 1(3):138-154.
673
674 El-Sayed AM, Mitchell VJ, Manning LM, Suckling DM (2011) New sex pheromone blend for the
675 Lightbrown Apple Moth, *Epiphyas postvittana*. *J Chem Ecol* **37**:640-646.
676
677 Ferkovich SM, Oliver JE, Dillard C (1982) Pheromone hydrolysis by cuticular and interior esterases of the
678 antennae, legs, and wings of the cabbage looper moth, *Trichoplusia ni* (Hübner). *J Chem Ecol* 8(5):859-
679 866.
680
681 Fox J & Weisberg S (2011). An R Companion to Applied Regression, Second Edition. Thousand Oaks CA:
682 Sage. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion
683
684 Frey JE & Bush GL (1990) *Rhagoletis* sibling species and host races differ in host odor recognition.
685 *Entomol Exp Appl* **57**:123-131.
686
687 Grula JW, McChesney JD, Taylor OR Jr (1980) Aphrodisiac pheromones of the sulfur butterflies *Colias*
688 *eurytheme* and *C. philodice* (Lepidoptera, Pieridae). *J Chem Ecol* **6**:241-256.
689
690 Ho HY, Tao YT, Tsai RS, Wu YL, Tseng HK, Chow YS (1996) Isolation, identification, and synthesis of sex
691 pheromone components of female tea cluster caterpillar, *Andraca bipunctata* Walker (Lepidoptera:
692 Bombycidae) in Taiwan. *J Chem Ecol* **22**(2):271-285.
693
694 Jacob S, McClintock MK, Zelano B, Ober C (2002) Paternally inherited HLA alleles are associated with
695 women's choice of male odor. *Nat Genet* **30**:175–179.

696 Jiggins CD (2017) The ecology and evolution of *Heliconius* butterflies. Oxford University Press (Oxford,
697 UK).

698 Jones, C.D. (2001) Extension of the Castle–Wright Effective Factor Estimator to Sex Linkage and
699 Haplodiploidy. *J Hered* **92**(3):274-276.

700 Kárpáti Z, Dekker T, Hansson BS (2007) Reversed functional topology in the antennal lobe of the male
701 European corn borer. *J Exp Biol* **211**:2841-2848.
702
703 Klein AL, de Araújo AM (2010) Courtship behavior of *Heliconius erato phyllis* (Lepidoptera, Nymphalidae)
704 towards virgin and mated females: conflict between attraction and repulsion signals? *J Ethology*
705 **28**(3):409-420.
706
707 Kochansky J, Cardé RT, Liebherr J, Roelofs WL (1975) Sex pheromone of the European corn borer,
708 *Ostrinia nubialis* (Lepidoptera: Pyralidae), in New York. *J Chem Ecol* **1**(2):225-231.

709 Koller M (2016) robustlmm: An R package for robust estimation of linear mixed-effects models. *J Stat*
710 *Software* **75**(6):1-24.
711
712 Koutroumpa FA, Kárpáti Z, Monsempes C, Hill SR, Hansson BS, Jacquin-Joly E, Krieger J, Dekker T (2014)
713 Shifts in sensory neuron identity parallel differences in pheromone preference in the European corn
714 borer. *Frontiers Ecol Evol* **2**:65.
715
716 Larsdotter-Mellstrom H, Eriksson K, Liblikas I I, Wiklund C, Borg-Karlson AK, Nylin S, Janz N, Carlsson MA
717 (2016) It's all in the mix: blend-specific behavioral response to a sexual pheromone in a butterfly.
718 *Frontiers Physiol* **7**:68.
719
720 Lassance J-M, Groot AT, Liénard MA, Antony B, Borgwardt C, Andersson F, Hedenström E, Heckel DG,
721 Löfstedt C (2010) Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex
722 pheromones. *Nature* **466**:486-489.
723
724 Lenth R (2018). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version
725 1.3.1. URL https://CRAN.R-project.org/package=emmeans
726
727 Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv
728 1303.3997.
729
730 Li H & Durbin R (2011) Inference of human population history from individual whole-genome sequences.
731 Nature 475:493-496.
732
733 Liénard MA, Wang H-L, Lassance J-M, Löfstedt C (2014) Sex pheromone biosynthetic pathways are
734 conserved between moths and the butterfly Bicyclus anynana. Nat Comms 5:3957.
735
736 Malausa T, Leniaud L, Martin J-F, Audiot P, Bourget D, Ponsard S, Lee S-F, Harrison RG, Dopman E (2007)
737 Molecular differentiation at nuclear loci in French host races of the European corn borer (*Ostrinia*
738 *nubilalis*). *Genetics* **176**(4):2343-2355.
739
740 Mann F, Vanjari S, Rosser N, Mann S, Dasmahapatra KK, Corbin C, Linares M, Pardo-Diaz C, Salazar C,
741 Jiggins C, Schulz S (2017) The scent chemistry of *Heliconius* wing androconia. *J Chem Ecol* **43**:843-857.
742

743 Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, Blaxter M, Manica A, Mallet
744 J, Jiggins CD (2013) Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies.
745 *Genome Res* **23**:1817-1828.
746
747 McCormick AC, Gershenzon J, Unsicker SB (2014) Little peaks with big effects: establishing the role of
748 minor plant volatiles in plant–insect interactions. *Plant, Cell & Environment* **37**(8):1836-1844.

749 McElfresh JS, Chen X, Ross DW, Millar JG (2000) Sex pheromone blend of the pandora moth
750 (Lepidoptera: Saturniidae), an outbreak pest in pine forests (Pinaceae). *Canadian Entomologist*
751 **132**(6):775-787.
752
753 Meinwald J, Meinwald YC, Mazzocchi PH (1969) Sex pheromone of the queen butterfly: chemistry.
754 *Science* **164**(3884):1174-1175.
755
756 Mérot C, Frérot B, Leppik E, Joron M (2015) Beyond magic traits: Multimodal mating cues in *Heliconius*
757 butterflies. *Evolution* **69**(11):2891-2904.
758
759 Merrill RM, Rastas P, Martin SH, Melo MC, Barker S, Davey J, McMillan WO, Jiggins CD (2019) Genetic
760 dissection of assortative mating behavior. *PLOS Biology* **17**(2):e2005902.
761
762 Richard M. Merrill, Bas Van Schooten, Janet A. Scott, and Chris D. Jiggins (2010) Pervasive genetic
763 associations between traits causing reproductive isolation in Heliconius butterflies. Proc R Soc Lond B
764 278(1705):511-518.
765
766 S. H. Montgomery & R. M. Merrill (2017) Divergence in brain composition during the early stages of
767 ecological specialization in Heliconius butterflies. J Evol Biol 30(3):571-582.
768
769 More, J. D., and Finney, N. S. 2002 A Simple and Advantageous Protocol for the Oxidation of Alcohols
770 with o-Iodoxybenzoic Acid (IBX). *Org Lett* **4**:3001–3003.
771
772 Nieberding CM, de Vos H, Schneider MV, Lassance J-M, Estramil N, Andersson J, Bång J, Hedenström E,
773 Löfstedt C, Brakefield PM (2008) The male sex pheromone of the butterfly *Bicyclus anynana*: towards an
774 evolutionary analysis. *PLOS ONE* **3**(7):e2751.
775
776 Nishida R, Schulz S, Kim CS, Fukami H, Kuwahara Y, Honda K, Hayashi N (1996) Male sex pheromone of a
777 giant danaine butterfly, *Idea leuconoe*. *J Chem Ecol* **22**(5):949-972.
778
779 Papke RS, Kemp DJ, Rutowski RL (2007) Multimodal signaling: structural ultraviolet reflectance predicts
780 male mating success better than pheromones in the butterfly *Colias eurytheme*. *Anim Behav* **73**:47-54.
781
782 Pessoa Pinharanda, A. L. (2017). The genomic basis of species barriers in Heliconius butterflies (Doctoral
783 thesis). https://doi.org/10.17863/CAM.16857
784
785 Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team (2018) nlme: Linear and Nonlinear Mixed Effects
786 Models. R package version 3.1-137, URL: https://CRAN.R-project.org/package=nlme.
787

Pires EV, Mendonça, Vaníčková L, Serra NSJ, da Silva RdCC, dos Santos DC, Campos RdS, Sant'Ana AEG, do Nascimento RR (2015) Identification and field and laboratory tests of the sex pheromone of *Cerconota anonella* Sepp. (Lepidoptera: Oecophoridae). *J Appl Entomol* **140**(1-2):72-80.

Pliske TE & Eisner T (1969) Sex pheromone of the queen butterfly: biology. *Science* **164**(3884):1170-1172.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rastas P (2017) Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. Bioinformatics 33(23):3726-3732.

Schulz S & Nishida R (1996) The pheromone system of the male danaine butterfly, *Idea leuconoe*. *Bioorg Med Chem* **4**(3):341-349.

Smadja CM & Butlin RK (2011) A framework for comparing processes of speciation in the presence of gene flow. Mol Ecol 20(24):5123-5140.

Stengl M (2010) Pheromone transduction in moths. *Frontiers Cell Neurosci* **4**:133.

Symonds MRE, Elgar MA (2007) The evolution of pheromone diversity. *Trends Ecol Evol* **23**(4):220-228.

Taga ME, Bassler BL (2003) Chemical communication among bacteria. *Proc Natl Acad Sci USA* **100** (suppl 2):14549-14554.

Tait C, Batra S, Ramaswamy SS, Feder JL, Olsson SB (2016) Sensory specificity and speciation: a potential neuronal pathway for host fruit odour discrimination in *Rhagoletis pomonella*. *Proc R Soc Lond B* **283**:20162101.

Tatsuki S, Kurihara M, Usui K, Ohguchi Y, Uchiumi K, Fukami J-i (1983) Sex pheromone of the rice stem borer, *Chilo suppressalis* (Walker) (Lepidoptera: Pyralidae): the third component, *Z*-9-hexadecenal. *Appl Entomol Zool* **18**(3):443-446.

Tumlinson JH, Brennan MM, Doolittle RE, Mitchell ER Brabham A, Mazomenos BE, Baumhover AH, Jackson DM (1989) Identification of a pheromone blend attractive to *Manduca sexta* (L.) males in a wind tunnel. *Arch Insect Biochem Physiol* **10**:255-271.

Steven M Van Belleghem, Pasi Rastas, Alexie Papanicolaou, Simon H Martin, Carlos F Arias, Megan A Supple, Joseph J Hanly, James Mallet, James J Lewis, Heather M Hines, Mayte Ruiz, Camilo Salazar, Mauricio Linares, Gilson RP Moreira, Chris D Jiggins, Brian A Counterman, W Owen McMillan, Riccardo Papa. (2017) Complex modular architecture around a simple toolkit of wing pattern genes. Nature Ecol Evol 1(3):0052.

Vane-Wright RI, Boppré M (1993) Visual and chemical signaling in butteflies: functional and phylogenetic perspectives. *Phil Trans R Soc Lond B* **340**:197-205.

Venables WN & Ripley BD (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

Via S (2012) Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. *Phil Trans R Soc Lond B* **367**(1587):451-460.

Wysocki CJ & Preti G (2004) Facts, fallacies, fears, and frustrations with human pheromones. *Anatomical Record Part A* **281A**(1):1201-1211.

Yang Z (2007) PAML 4: a program package for phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**: 1586-1591

Yew JY & Chung H (2015) Insect pheromones: An overview of function, form, and discovery. Prog Lipid Res 59:88-105.

Yildizhan S, van Loon J, Sramkova A, Ayasse M, Arsene C, ten Broeke C, Schulz S (2009) Aphrodisiac pheromones from the wings of the Small Cabbage White and Large Cabbage White butterflies, *Pieris rapae* and *Pieris brassicae*. *ChemBioChem* **10**(10):1666-1677.

Zufall F & Leinders-Zufall T (2000) The cellular and molecular basis of odor adaptation. *Chem Senses* **25**:473-481.