Evolution of chemical signals in *Heliconius* butterflies

Bruna Cama

Doctor of Philosophy

University of York
Biology
January 2022

Abstract

Speciation requires a reduction in gene flow between species. This can be achieved even in sympatry if reproductive barrier traits are under divergent ecological or sexual selection. Chemical signals are the oldest form of communication in the living world, and have promoted speciation in many taxa through the process of "sensory drive". *Heliconius* butterflies are model organisms in speciation, but few studies so far have been carried out on the evolution, genetics, physiology and regulation of their pheromones. In this study, I focus on two types of *Heliconius* pheromones: male sex pheromones (MSPs) involved in courtship, and antiaphrodisiacs used to enforce female monogamy. I investigated the macroevolution of MSPs across the Heliconiini tribe and found a burst of diversification at the base of the most speciose clade, and also that MSP divergence between sister species is strongly affected by the extent of sympatry. Both patterns are consistent with MSPs role as reproductive barriers in these butterflies. Using genetic crosses between two sympatric sister species, *H. elevatus* and *H. pardalinus*, I then mapped QTLs controlling differences in seven putative MSPs, and found that most QTLs were clustered and found in areas of genomic divergence between the species. For comparison with MSPs, I investigated the macroevolution of antiaphrodisiacs. I did not observe any evidence for a role of antiaphrodisiacs in reproductive isolation, but I identified some candidate compounds that may be behaviourally active, and observed an effect of host plant on antiaphrodisiac composition. Lastly, I found that the MSP and antiaphrodisiac blend compositions in *H. charithonia* and *H. atthis*, species with different mating strategies, to be strongly age-dependent. Despite these advances, the complexity of chemical signalling in these butterflies remains largely unknown. The candidate compounds and genes that I have uncovered are targets of future research through detailed behavioural assays, and genetic manipulation.

List of Contents

Acknowledgements

The years of my PhD have been a transformative time for me, as well as for the world around me, and things have not always been easy. But as a result, the support I received from everyone around me felt more real than ever. As I sit now putting the finishing touches on my thesis, I would like to thank all the people that were involved in this project and that were around for me during this time, making the past four years an unforgettable and invaluable experience.

Before moving on to the wonderful people that were part of this project, I'd like to acknowledge my funding sources. My project was funded by a Natural Environment Research Council (NERC) studentship as part of the Accepting the Challenges of a Changing Environment (ACCE) program, and by our case partner, the Smithsonian Tropical Research Institute (STRI). I also thank the University of York and the Department of Biology for this opportunity, and for being my second home for so long.

I would like to first and foremost thank my primary supervisor, Kanchon Dasmahapatra, for being a constantly encouraging and supportive presence through good and bad times. He provided me with training, guidance and feedback and was always there for me not only through my PhD, but through my MSc by Research as well, to brainstorm, to give advice, or sometimes simply to let me vent about whatever problems I'd come across. I am also extremely grateful to Kanchon for encouraging me to undertake my first fieldwork, which I consider the most positively lifechanging experience I've ever had.

I thank my secondary supervisor, Jane Thomas-Oates, for her incredible enthusiasm in my research topic, and for her support and her advice on the more technical aspects of my work, always being able to bring a fresh and helpful (and Chemistry-based!) perspective to the table in our discussions. Having both Kanchon and her as supervisors has been a wonderfully complementary experience.

I thank the members of Kanchon's lab group both present and past: Jake Morris, Josie Monaghan, Michael Lawson, Melissa Minter, Edd Page, Alex Payne and Lewis White, for being always friendly and supportive, for the company inside and outside of work, and for making my time here as pleasant as can be, even though most of our time together had to be spent remotely. I also thank Neil Rosser from the University of Harvard for his help and advice on the genomic side of my research, as well as for generating a large part of the genomic data I used in my analysis. I especially thank Lucie Queste, with whom I shared my labwork and my field season, for her positivity, friendship and support as we lived and worked together.

I would also like to thank Jane Thomas-Oates' lab and collaborators: Brett Sallach, Ed Bergström, Lea Drieu, John Wilkinson, Isabella Gaffney, Angelo Lopez, Alice Marshall, Ben West, Simeon Hubbard,

Hannah Spencer, Rebecca Morgan, and Fazeelah Munir, as being able to share my PhD with such a diverse and lively group has always been both interesting and fun. I also thank and Karl Heaton at the University of York GCMS center, for his help and support in using the York instruments when COVID made it necessary. My last chapter would be entirely missing had he and Jane not stepped in to help.

I would then like to thank Stefan Schulz, who has been my supervisor during my visits to Braunschweig TU, as well as his group, in particular Stephanie Ehelrs and Daiane Szczerbowski whom shared both their data and their expertise on the matter of GCMS. This entire piece of research would not exist without Prof. Schulz, his technical guidance and his sheer volume of knowledge on the topic of chemical ecology always being enlightening.

Of course, I also wish to thank my on-field STRI supervisor Owen McMillan, whom I first met at the *Heliconius* Meeting when I was still a MSc student, and whom I remember remarking "We gotta get you down to Panama" after hearing my presentation that year. He helped me achieve that at each step along the way, and working with the Smithsonian Tropical Research Institute has been both an honour and a great pleasure. I will never forget my time there, nor the Gamboa crew I had the pleasure of meeting and collaborating with: Kathy Darragh, Kelsey Byers, Sebastian Mena, Rémi Mauxion, Lina Melo, Oscar Paneso, Diana Abondano, Sylvia Garza, Chi-Yun Kuo, Monica Monllor, Edgardo Santiago, Laura Hebberecht, Fletcher Young, Jamie Musgrove, Nate Edelman and Luca Livraghi, as well as all of the other STRI crew researchers that have come and gone. And speaking of collaborators, I also have to thank Nicola Nadeau and her lab at the University of Sheffield for aiding me in my data collection for my antiaphrodisiac chapter.

I have some really important people to thank outside of my research too. First, I thank my friends throughout the world: my friends in Italy both old and new, whom I knew I could always come back to when things got rough, my undergraduate pals in the UK, and my friends in California, with whom I've gone on many adventures, and who knows where I would be without them. Lastly, I want to thank my whole family, but especially mom and dad, and my little siblings Simona and Vittorio. Their support has been unending and unwavering, and their sheer strength and love have always inspired me to keep trying, in the world research and outside.

I dedicate this work to my little brother Riccardo, who left us way too soon and shortly before the beginning of my PhD. A wise, joyful and hardworking soul, a tremendously intelligent man, and an unwavering inspiration, in research and in life.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Most of the data presented in this thesis was collected, processed and analyzed by me. Nonetheless, the work presented here has been a collaborative effort, and my other contributors are as follows:

Samples from Ecuadorian species for Chapter 1 and 3 were collected and processed by Stephanie Ehelrs and Daiane Szczerbowski, under the supervision of Stefan Schulz (Braunschweig Technische Universität) and Chris Jiggins (University of Cambridge). Stephanie Ehlers also identified a number of unknown compounds, and the Schulz laboratory collectively compiled most of the libraries I used for compound identification. H. timareta and H. eleuchia samples were collected and processed by Kathy Darragh, under the supervision of Chris Jiggins. While these collaborators carried out their own analyses on their collected data, all analyses presented in this study are my own, and I designed all of the experiments.

All RAD libraries used for Chapter 2 were prepared by myself, Lucie Queste, Michaela Nelson and Kanchon Dasmahapatra at the University of York, and linkage maps were generated by Kanchon Dasmahapatra and Neil Rosser (Harvard University). Neil Rosser also oversaw and offered advice on the QTL analysis, as well as preparing the script used to extract LOD scores, confidence intervals and genotype:phenotype relationships for each compound from F2 crosses, which I then tweaked to suit my analysis. The pure species data I present in Chapter 2 is published in:

Rosser, N., Queste, L., Cama, B. et al. (2019) 'Geographic contrasts between pre- and postzygotic barriers are consistent with reinforcement in *Heliconius* butterflies', Evolution, 73(9), pp. 1821–1838. doi: 10.1111/evo.13804.

Heliconius sara for Chapter 4 were reared at the University of Sheffield by Nicola Nadeau, Eve Scott and Harriet O'Roarty, who also carried out the tissue extractions.

General Introduction

Evolution, speciation and communication systems

The genic view of speciation

Understanding the processes by which new species form is a central topic in evolutionary biology. New species arise as what was once a single population diverges into two ecologically distinct and isolated groups (Coyne and Orr, 2004), a process known as speciation. The concept of species, and accordingly, theories as to how new species arise, have in themselves evolved over time. According to the biological species concept (BSC) (Mayr, 1942) a species is a reproductively isolated unit wherein conspecific individuals can interbreed, but cannot cross with heterospecific individuals, in turn making lack of gene flow a prerequisite of biological species. Under this definition, underlying the separation of two biological species is widespread genomic divergence, which is achieved via co-adaptation of many genes at the same time, so that the whole genome operates as a unit rather than a set not fully co-dependent genes (Mayr, 1963). Thus, under the BSC, changes in environment that trigger speciation simultaneously affect many regions of the genome at once, with the result being the emergence of new traits.

As time passed from the original definition of the BSC, the concept of species and the process of speciation have become open to new perspectives (see for example: Mishler and Theriot, 2000; Wheeler and Rudolf, 2000; Wiley and Mayden, 2000), and nowadays, the BSC is mostly considered one extreme end of the so-called "speciation continuum" (Shaw and Mullen, 2014). The speciation continuum represents the full range of intermediate scenarios going from a single population where all individuals can interbreed freely, to completely reproductively isolated biological species, as genetic and phenotypic changes accumulate leading to progressively more divergent populations, until the two genomes are fully incompatible (Shaw and Mullen, 2014). With this new, flexible view of the process of population divergence, coupled with new molecular insights, the scientific community moved towards a new concept of species: the genic view of speciation (Wu, 2001; Wu and Ting, 2004), where the focus shifts away from reproductive isolation and towards genetic isolation instead.

Under the genic view of species, a complete lack of gene flow is not a prerequisite for speciation. By this view, speciation occurs when two emergent species will be able to continue diverging, and will remain distinct, even in the face of gene flow (Wu, 2001). The genome does not diverge uniformly, but rather, divergence happens at select loci wherein genes that drive reproductive isolation termed "speciation genes" are found (Wu and Ting, 2004; Nosil and Schluter, 2011). Loci that are not involved in speciation, on the other hand, may continue to be shared and be the subject of admixture and gene

flow between the emerging species, in the case that interaction between them is possible (e.g. if they have overlapping ranges), making divergence heterogenous across the genome (Nosil and Feder, 2012). Speciation genes drive speciation by affecting mate preference, ecological adaptation, and making hybrids less fit than the pure species, (Wu and Ting, 2004; Nosil and Schluter, 2011). In this way, the single genes promote reproductive isolation, rather than physical separation between the emerging species (Rieseberg and Blackman, 2010; Nosil and Schluter, 2011). It follows that under the genic view, speciation is possible even in scenarios of sympatry (full range overlap) or parapatry (partial range overlap) between emerging species, as long as divergent selection is acting on speciation genes.

Ecological speciation, reproductive barriers and "speciation genes"

Divergent selection is at the very core of ecological speciation, as the process that sets it apart from other speciation models driven by random events (Rundle and Nosil, 2005), such as mutations, genetic drift, population bottleneck events and genome duplications among others (Rundle and Nosil, 2005; Sobel *et al.*, 2010). Under ecological speciation, divergence is achieved via adaptation to different niches within the environment, which can be dictated by differential resource availability, climate, levels of predation or competition and other features of the habitat (Schluter, 2001; Rundle and Nosil, 2005). Speciation genes are the link between divergent selection dictated by these differences in ecology, and reproductive isolation (Rundle and Nosil, 2005).

By promoting reproductive isolation under divergent selection, speciation genes produce reproductive barriers. A reproductive barrier is any behavioural, physiological, or otherwise ecological aspect of a species, or any combination of the above, that contributes to reducing gene flow with other species. Reproductive barriers can take effect before mating and prior to the fusion of the parental gametes (prezygotic barriers) or after (postzygotic barriers), and in fact reproductive isolation can be achieved through an accumulation of both types of barrier (Naisbit, Jiggins and Mallet, 2001; Sobel *et al.*, 2010; Rosser *et al.*, 2019). Thus, speciation genes can cover a great amount of different functions while having an isolating effect, and identifying speciation genes has been a major point of interest in speciation research (Sobel *et al.*, 2010; Nosil and Schluter, 2011).

Examples of speciation genes have been discovered in cases of both prezygotic and postzygotic isolation. Postzygotic isolation acts on the hybrid itself, in a number of ways (Orr, 1995). In the most extreme cases, postzygotic speciation genes can make the hybrid inviable, as is the case in *Drosophila melanogaster* and its sister species, wherein a gene called *Hmr* (*Hybrid male rescue*) is linked to hybrid lethality as it encodes a vital transcription factor whose function is disrupted in hybridization (Barbash *et al.*, 2003). This gene was found to evolve at a very fast rate in *Drosophila* species, with widespread

changes in its nucleotide sequence even among closely related sister species, which led Barbash et al. to theorize that rapid evolution may be common among speciation genes (Barbash *et al.*, 2003; Barbash, Awadalla and Tarone, 2004). But postzygotic isolation is not always so extreme: in other cases, it can lead to infertile, but otherwise viable hybrids. Within *Drosophila,* male crosses between *D. simulans* and *D. mauritiana* are sterile, and in this case the sterility maps to a different gene called *Ods,* which like *Hmr* carried signs of positive selection and rapid recent evolution (Ting *et al.*, 1998). In other cases, even if the two species' genomes are compatible, selection against hybridzation is extrinsic rather than instrinsic, with hybrids being negatively affected by ecological factors in a way the pure species are not. For example, hybrids between two species of sticklebacks adapted to two different environments show intermediate traits compared to the two parental species, and grow more slowly than either parental species in either of the two environment, being more poorly adapted to both lifestyles (Hatfield and Schluter, 1999). In *Heliconius* butterflies, known for Müllerian mimicry of their warning colour pattern which signals their unpalatability to predators, hybrid individuals that show "imperfect" intermediate patterns are selected against, as predators do not recognize the arrangement of their spots as a warning sign (Jiggins *et al.*, 2001; Naisbit, Jiggins and Mallet, 2001). However, the *Heliconius* colour pattern also acts as a prezygotic barrier, as the butterflies show assortative mating towards individuals bearing the same colour pattern as themselves, and in the closely related *H. melpomene* and *H. cydno,* the loci for colour pattern and mate preference are actually genetically linked (Jiggins *et al.*, 2001). Such genetic linkage, or when a trait that simultaneously affects adaptation and mate choice, can greatly facilitate speciation in the absence of physical separation between the diverging populations.

Prezygotic barriers take effect before fusion of the gametes, often before mating. As ecological speciation can occur in allopatry as well as in sympatry (Schluter, 2001), a geographical feature separating the diverging populations would be a prezygotic reproductive barrier, but while geographical separation often affects the whole genome due to the accumulation of random changes, speciation genes truly come into play in situations of sympatry or parapatry (Wu and Ting, 2004; Rundle and Nosil, 2005; Nosil and Schluter, 2011). In these contexts, there are many kinds of prezygotic reproductive barriers, encompassing any ecological phenomenon that creates a partitioning between subpopulations so that interbreeding becomes less likely and gene flow is reduced. For example, in *Mimulus guttatus* monkeyflowers, postzygotic isolation is absent, yet prezygotic isolation between varieties is dictated by different flowering times (Lowry, Rockwood and Willis, 2008). In plants in general, shifts in flower colour dictated by losses or changes in function of genes often lead to pollinator shifts, which in turn lead to reproductive isolation (Quattrocchio *et al.*, 1999; Schwinn *et al.*, 2006; Rieseberg and Blackman, 2010). Different closely related species of cichlids

inhabiting the African rift lakes are known to mate assortatively based on specific colour patterns or hues (Seehausen and Van Alphen, 1998; Genner *et al.*, 2007). In sympatric field crickets *Gryllus texensis* and *Gryllus rubens,* slight differences in mating call are genetically linked to mate preference as seen in *Heliconius* (Jiggins *et al.*, 2001), as females prefer calls of their own species, and thus although the hybrids are viable, species identity is preserved (Gray and Cade, 2000). In Darwin's finches of the *Geospiza* genus, different feeding behaviours lead to altered beak and body morphology, which in turn affects their song patterns, utilized by females to discern and mate with conspecifics (Ratcliffe and Grant, 1985; Podos, 2001). When considering these examples of prezygotic pre-mating reproductive barriers, a common property emerges among not all, but many of them: being involved in communication.

How signals shape evolution via sexual selection

The diversity of signals and signalling behaviours observed across the tree of life is staggering, and communication systems contribute significantly to speciation and evolution in general (Endler, 1992). The environment wherein signalling takes place shapes the sensory perception systems of an organism, which in turn shape the type of signals utilized to convey information and the physiological features needed to produce them, as well as the behaviours needed to transfer them to other individuals. The signal recipients, on the other hand, evolve to decode the information carried by the signal, be they conspecifics using it for mate choice or even predators wanting to hijack it (Endler, 1992). With so many interactions between different factors, communication systems can drive the evolution of taxa in a certain direction, a phenomenon known as "sensory drive" (Endler, 1992; Boughman, 2002). Divergent selection can occur via sensory drive, and in turn sensory drive can affect reproductive isolation (Boughman, 2002). In the following section, I will cover how the concepts of sensory drive, divergent selection, mate choice and reproductive isolation can be tightly intertwined when discussing ecological speciation.

As communication is so dependent on the organism's ecology, changes in habitat between two species can drive divergence between signalling traits, and if the signal itself is used in mate choice, it is likely to become a reproductive barrier (Gray and Cade, 2000; Boughman, 2002; Masta and Maddison, 2002). In sexual selection, mate choice creates disparity in mating success among members of the species based on preference for specific sexual traits. If different subpopulations favour different characteristics, this can lead to prezygotic isolation (Panhuis *et al.*, 2001). There are several possible reasons why disparity in sexual signalling traits and preferences might be introduced in the first place. It can occur as an indirect result of morphological changes related to other aspects of an organism's ecology (e.g. changes in body size in birds leading to different throat conformations, which affect the song (Podos, 2001)), or with the purpose of minimizing ambiguity in species recognition, or as a result

of adaptation to different signalling environments created by differences in habitat, or any combination of the above (Seddon, 2005). It follows that divergent ecological selection and sexual selection can be tightly interconnected to each other, especially when they are genetically linked to loci controlling mate preference.

An example of this is seen in sympatric three-spined sticklebacks inhabiting lakes in British Columbia, where divergent populations have formed in habitats with differing light environments as females adapted to different light conditions favour different male nuptial colorations based on the colour range they can perceive (Boughman, 2001). A similar phenomenon has been demonstrated in sympatric species of Lake Victoria cichlids that favour different light levels (Seehausen *et al.*, 2008). In both of these cases, male nuptial colour is the signal, and it is both under divergent selection and involved in mate choice. The signal does not need to be visual for it to be affected by the environment. For example, In Neotropical antbirds, forest habitat affects song pitch as species that live in the understory or other densely vegetated environments produce higher-pitched songs than species that favour more open habitats (Seddon, 2005). One further important result of the antbird study is that while the songs were generally species-specific, of the many featured species, sympatric closely related species pairs showed greater song divergence than allopatric pairs (Seddon, 2005). This is a typical property of reproductive barrier traits known as reproductive character displacement, and it is an evolutionary response to sympatry between close relatives, with the aim of minimizing disadvantageous interactions with heterospecifics (Pfennig and Pfennig, 2009). Reproductive character displacement further enhances reproductive isolation and is also known as “reinforcement” in cases where it stems from selection to prevent hybridization (Pfennig and Rice, 2014).

Signalling systems under sexual selection can even potentially contribute to evolutionary radiations. While divergence in signalling traits can occur after the population has already begun to diverge (Arnegard *et al.*, 2010), there are known cases of evolutionary radiations where sexual selection on the signalling traits themselves promoted speciation (Masta and Maddison, 2002; Mendelson and Shaw, 2005; Boul *et al.*, 2007; Arnegard *et al.*, 2010). This can occur especially when an innovation in the signalling system opens up new niches represented by unexploited signal space, and it can result in rapid diversification of the signal, which can easily become differentiated even among closely related species pairs (Arnegard *et al.*, 2010)- in other words, the signal becomes evolutionarily labile, and not phylogenetically constrained. As a peculiar example, Mormyridae electric fish utilize electric pulses as part of their courtship, a very uncommon type of communication that is thus neither exploited or eavesdropped on by other species, nor easily confounded in their environment (Arnegard *et al.*, 2010). This signal is what mate choice is entirely based on in this fish radiation and is strongly differentiated between both closely related and distantly related species, showing little to no

phylogenetic constraint, whereas ecological traits differentiated in a less explosive way and reflect taxonomic relationships between species (Arnegard *et al.*, 2010). An even more extreme example is seen in *Laupala* Hawaiian swordtail crickets, which have undergone (and are currently undergoing) one of the most rapid evolutionary radiations in the animal kingdom (Shaw and Lesnick, 2009) yet are ecologically and morphologically indistinct for the most part, except for their extremely diverse mating calls (Mendelson and Shaw, 2005), controlled by genes genetically linked to female choice loci in at least one species pair (Shaw and Lesnick, 2009; Wiley and Shaw, 2010).

Having established the ways in which communication can be affected by divergent selection both ecological and sexual, and how in turn it can have a key role in promoting species divergence, it is now important to discuss one specific form of communication: chemical signals.

Chemical signals in the animal kingdom

In this section, I will discuss chemical signals first from an ecological and physiological standpoint, and then in the context of evolution and speciation. Chemosensation is the ability of an organism to perceive molecules from its environment and react to them either directly or via activation of neural pathways, and is one of the most ancient and most widespread sensory modes in the living world (Hildebrand and Shepherd, 1997). Most, if not all lifeforms have chemosensory systems which play essential roles in their survival, and in vertebrates, chemosensation fully encompasses such familiar "senses" as olfaction and gustation (Hildebrand and Shepherd, 1997). Chemical perception is used to detect and interact with other individuals, conspecific or not, to find food, and to discern and avoid harmful substances. It is also essential for multicellularity, as it is how cells interact and respond to each other. It should come as no surprise, then, that chemical signals are one of the most ubiquitous forms of communication across all living organisms.

Chemical signals: one of the most widespread forms of communication

Most organisms, ranging from bacteria to unicellular eukaryotes, and from plants to animals, utilize some form of chemical communication (Ginzel, 2009). The latter often takes the form of semiochemicals: biologically active substances that produce a behavioural response in receiving individuals (Martin and Hine, 2008; Ginzel, 2009). Semiochemicals are known as allelochemicals when used to communicate across species, and as pheromones when used to communicate within species (Karlson and Lüscher, 1959; Ginzel, 2009). The composition of these substances is extremely variable. They can be made out of a single component or several, and the components themselves can range in weight and complexity, from relatively simple and small molecules like formic acid, part of ant alarm pheromones and containing a single carbon nucleus (Ayre and Blum, 1971; Bradshaw, Baker and Howse, 1975; Löfqvist, 1976), to nucleotides, observed in crustaceans (Fletcher *et al.*, 2021), through

to large and complex like the proteins that make up some mammalian pheromones (Hurst *et al.*, 2001; Chamero *et al.*, 2007). Most of the topics I will discuss here apply similarly to both allelochemicals and pheromones, in terms of their composition, biosynthesis and the many functions they can cover, with the only meaningful difference being the receiving party. However, I will be focusing exclusively on pheromones.

Pheromones and their functions

Pheromones are behaviourally active compounds or substances, used for communication between conspecifics. They are widespread across the tree of life, being found in bacteria (Taga and Bassler, 2003; Williams *et al.*, 2007), fungi (Casselton and Olesnicky, 1998; Casselton, 2002; Nieuwenhuis and Aanen, 2012), algae (Gassmann, Müller and Fritz, 1995; Frenkel, Vyverman and Pohnert, 2014), and in many different animal taxa, including but not limited to annelids (Hardege *et al.*, 1998; Watson, Hamilton and Tuffnail, 2005), rotifers (Snell, 1998), mollusks (Cimino *et al.*, 1991; Painter *et al.*, 2004), vertebrates (Gosling and Roberts, 2001; Schaal *et al.*, 2003; Chamero *et al.*, 2007), crustaceans (Fletcher *et al.*, 2021) and insects (Regnier and Law, 1968; Howard and Blomquist, 2005; Yew and Chung, 2015).

Pheromones can take on a wide range of functions. Because their composition is species-specific, species recognition is one of their most basic purposes (Howard and Blomquist, 2005). One of the most primitive forms of chemical communication, for example, is quorum sensing in prokaryotes, wherein bacterial cells are able to scan the environment for pheromone molecules emitted by other bacteria, assess the presence or absence and abundance of conspecific individuals or potential competitors, and react accordingly, in a very basic display of "social" behaviour (Taga and Bassler, 2003; Jacob *et al.*, 2004; Williams *et al.*, 2007). Other types of pheromones mediate collaboration between conspecifics. Aggregation pheromones attract individuals of the same species towards a certain point, and are particularly common in species that exploit patchy resources and benefit from being able to locate them easily (Matthew R.E. Symonds and Elgar, 2004; Ginzel, 2009; M.R.E. Symonds and Gitau-Clarke, 2016; Jurenka *et al.*, 2017). Trail or recruitment pheromones are similar as they prompt conspecifics to move towards a specific direction, however in this case the behaviourally active chemicals form a lane for the individual to follow (Regnier and Law, 1968; Ginzel, 2009; Yew and Chung, 2015). These are especially in common and important in ants, as workers lay down these pheromone trails to signal the fastest route to a resource or migration path (Ayre and Blum, 1971; Couzin and Franks, 2003; Yew and Chung, 2015), albeit it is found in other taxa too (Wilmes *et al.*, 2012; Ng *et al.*, 2013). Alarm pheromones activate defence behaviour in social animals, generally inducing a state of alertness that intensifies as the pheromone concentration increases, until it reaches a threshold past which the colony or group enters a frenzied attack phase, occasionally directed at the

source of the alarm pheromone itself (Regnier and Law, 1968). Alarm pheromones are common among social insects (Moser, Brownlee and Silverstein, 1968; Regnier and Law, 1968; Boch and Rothenbuhler, 1974; Bradshaw, Baker and Howse, 1975; Löfqvist, 1976; Ginzel, 2009; Moritz and Bürgin, 2010), and are in general most beneficial for species that form social communities.

These are only some of the many behaviours pheromones can affect. It is also possible for pheromones to cover more than one function at the same time. Moreover, it is common for organisms to produce more than one type of pheromone for different occasions. One type of pheromones, however, has received particular interest in the field of chemical ecology as well as in the field of evolution: sex pheromones.

Sex pheromones: diversity and discovery

Sex pheromones are diverse in form, function and complexity (Johansson and Jones, 2007; Yew and Chung, 2015). An individual seeking to mate has to convey multiple aspects of its status to its prospective partner: its species, sex, reproductive status, and in some cases even its age and quality as a mate (Johansson and Jones, 2007). Some sex pheromones have aphrodisiac function, acting as sexual stimulants (Grula, McChesney and Taylor, 1980; Ferveur, 2005; Nieberding *et al.*, 2008; Darragh *et al.*, 2017), while others have antiaphrodisiac function, discouraging mating (Andersson *et al*, 2000; Andersson *et al.*, 2004; Brent and Byers, 2011; Estrada *et al.*, 2011), and others still, like bombykol, are simply attractants for individuals of the opposite sex (Yew and Chung, 2015). The site of emission for sex pheromones varies too: in Diptera for example, sexual signals are transmitted via cuticular hydrocarbons and are thus contact-based (Carlson *et al.*, 1971; Ferveur, 2005; Howard and Blomquist, 2005; Wicker-Thomas, 2007, 2011; Shahandeh, Pischedda and Turner, 2018), whereas in Lepidoptera they are released by specialized structures on the wings known as androconia, or from abdominal or wing hair pencils, and are generally active over long ranges (Schulz, Boppre and Vane-Wright, 1993; Nishida *et al.*, 1996; Ômura and Honda, 2011; Wang *et al.*, 2014; Vanjari *et al.*, 2015; Darragh *et al.*, 2017).

The first sex pheromone, and the first pheromone overall to be chemically characterised, was discovered in the silk moth *Bombyx mori* in 1959 (Butenandt, Beckmann and Hecker, 1961), and was named bombykol. This relatively simple compound, also named (10E,12Z)-hexadeca-10,12-dien-1-ol, incidentally sums up many of the characteristics commonly observed in pheromone compounds, particularly in insects. In terms of its biosynthesis, it is an alcohol obtained from a chain of reactions starting from one main precursor, a fatty acid (FA) in its fatty acyl-CoA form. It is desaturated, bearing two double bonds that are introduced by an enzyme known as a desaturase (Jurenka, 2004; Howard and Blomquist, 2005), in particular a fatty acyl Δ11-desaturase (Jurenka, 2004; Moto *et al.*, 2004). The

product is then transformed into the final alcohol compound (Moto *et al.*, 2003), by another enzyme, a fatty acyl reductase. This setup is common in insect sex pheromones, as their behaviourally active compounds are often (though not exclusively) fatty acid derivatives that can undergo modifications such as elongation, desaturation and reduction (Jurenka and Roelofs, 1993; Tillman *et al.*, 1999; Jurenka, 2004; Howard and Blomquist, 2005; Yew and Chung, 2015; Jurenka *et al.*, 2017), sometimes accompanied by host plant-derived products (Jurenka, 2004; Howard and Blomquist, 2005; Yew and Chung, 2015). In terms of function, bombykol is a long distance mating pheromone female silk moths use to attract males (Butenandt, Beckmann and Hecker, 1961). Since its discovery, sex pheromones have become the most studied category of chemical signals, particularly in Lepidoptera and Diptera (Regnier and Law, 1968; Johansson and Jones, 2007; Ginzel, 2009).

Not only are Lepidoptera the organisms wherein the first sex pheromone was discovered and characterized (Butenandt, Beckmann and Hecker, 1961), kickstarting decades of research in the field of chemical ecology, they also took a central role in research on the topic (Symonds and Elgar, 2008), and several species have had their chemical signatures fully characterized. Moths have received more attention than butterflies due to their greater reliance on olfactory signals for mating; as moths are nocturnal, males cannot rely on sight to locate females, therefore female moths have evolved long-distance pheromone that they release in plumes to call for the males (Hansson, 1995; De Bruyne and Baker, 2008; Allison and Carde, 2017). Correspondingly, males have evolved specialized structures on their antennae packed with olfactory receptors to be able to detect even a few molecules from these plumes from a long distance, hence the extravagant feathery antenna shapes seen in some moth species (Hansson, 1995; Allison and Carde, 2017). Furthermore, as many moths are considered agricultural pests, their sex pheromones have received considerable attention as pest control measures, serving as olfactory traps for males (Carde and Minks, 1995; Witzgall, Kirsch and Cork, 2010). Butterflies, meanwhile, are known to be vision-oriented organisms; their more acute vision, diurnal lifestyle and bright, species-specific colour patterns imply that visual signals take precedence over chemical signals when it comes to locating a mate (Vane-Wright and Boppre, 1993; Hansson, 1995; Liénard *et al.*, 2014), and so their pheromones have received comparatively less attention. This is also reflected by their less elaborate antennae (Hansson, 1995).

Nonetheless, butterflies also utilize pheromones and have species-specific chemical signatures, though with some key differences from moths. In butterflies, females are not known to perform long-distance “calling” with their pheromone plumes, and pheromone emissions are mostly performed by the males instead, in a reversal of sexual signalling roles (Liénard *et al.*, 2014). Rather than calling females, male butterflies visually locate them and then use their wing androconia to release a bouquet of chemical signals during courtship, which contains enough information for the receiving females to

choose whether to accept or reject them as mates (Lundgren and Bergström, 1975; Nieberding *et al.*, 2008; Heuskin *et al.*, 2014; Darragh *et al.*, 2017; Li *et al.*, 2017). How females decode this information to make their choice is currently largely unknown, and it must be noted than in both moths and butterflies respectively, males and females are also likely to emit their own short-distance chemical signals, but very little is known about them (Liénard *et al.*, 2014).

In terms of composition, moth and butterfly sex pheromones both feature multiple types of fatty acid derivatives, including but not limited to alcohols, aldehydes, alkenes, alkanes and esters. It has been shown that the fatty acid metabolic pathway used to produce bombykol, is conserved between moths and butterflies (Liénard *et al.*, 2014), and it involves any combination of fatty acyl-CoA elongation, desaturation, reduction and oxidation reactions performed by a widely conserved set of enzymes (Reed *et al.*, 1994, 1995; Tillman *et al.*, 1999; Wicker-Thomas, 2011; Liénard *et al.*, 2014; He *et al.*, 2017; Mann *et al.*, 2017). Butterfly pheromones however tend to be more complex than moth pheromones, and can include, other than fatty acid derivatives, a wider range of different products (Jurenka *et al.*, 2017) among which are terpenoids (Hayashi *et al.*, 1985; Johansson and Jones, 2007; Schulz *et al.*, 2008; Ehlers *et al.*, 2021), nitrogen-rich compounds (Schulz, Boppre and Vane-Wright, 1993; Nishida *et al.*, 1996) and aromatic compounds (Nishida *et al.*, 1996; Mozuraitis *et al.*, 2019), as well as other types of plant derivatives, which can be produced *de-novo* by the butterflies (Darragh, Orteu, *et al.*, 2019) or transformed from host plant compounds (Schulz, Yildizhan and van Loon, 2011). Despite the great variability of their composition in both butterflies and moths, receiving individuals are able to discern their conspecifics' chemical signatures from even very similar ones based on minute differences (Cardé *et al.*, 1977; Glover, Tang and Roelofs, 1987; Bengtsson *et al.*, 2014).

Pheromones in speciation

The composition of sex pheromones, their behavioural effect and their practical applications have been abundantly researched in many different species, yet the role of sex pheromones in speciation remains poorly explored compared to other, more easily perceptible traits. But as sex pheromones are involved in species recognition and mate choice, and especially given the ability of individuals to discriminate against heterospecific blends, they are ideal candidates for promoting reproductive isolation. A small change in pheromone composition, which may even be achieved through mutations at a single locus, could part the population in two, between individuals that preferentially mate with the original blend and those that favour the modified blend, and trigger the speciation process (Smadja and Butlin, 2009; Wicker-Thomas, 2011). Furthermore, changes in pheromone composition would necessarily be linked to behavioural and physiological changes needed to perceive and react to the new blend, potentially accelerating the process of population divergence (Smadja and Butlin, 2009).

A very classic and well-understood case of pheromone divergence mediating reproductive isolation is that of the European corn borer moth *Ostrinia nubilalis.* This moth's pheromones consist of a mixture of two forms of the same fatty acid derivative, (Z)-11-tetradecenyl acetate and (E)-11-tetradecenyl acetate, which differ only by the orientation of their Δ11 double bond (Kochansky *et al.*, 1975). In the wild, *O. nubilalis* is polymorphic, and exists in two distinct non-interbreeding populations that are ecologically very similar and have largely non divergent genomes, but females produce the two compounds in different ratios: the Z-type produces a 97:3 ratio of (Z)-11-tetradecenyl acetate to (E)-11-tetradecenyl acetate, while this ratio is almost exactly reversed in the E-type (Glover, Tang and Roelofs, 1987; Peña *et al.*, 1988). Z-type males are only attracted to females of their own type and vice versa, and hybrids in the wild are very rare (Wicker-Thomas, 2011), meaning this difference in female pheromones is the basis of reproductive isolation. There are many other known cases of behavioural isolation, paired with a lack of gene flow, caused primarily by pheromone differentiation (Smadja and Butlin, 2009). Many examples are found in moths (Löfstedt, Herrebout and Menken, 1991; Foster *et al.*, 1997; Monti *et al.*, 1997; Groot *et al.*, 2014), but it is a widespread phenomenon across animal taxa (reviewed in Smadja and Butlin, 2009).

In most cases, the divergence in pheromone composition affects the ratios of different types of compounds rather than their presence or absence (Smadja and Butlin, 2009), but there are exceptions where the main difference is given by lack of blend components in one of the two populations (McElfresh and Millar, 2001), and cases of sympatric sister species pairs with strongly differentiated pheromone blends do exist (Zimmermann, Ramirez and Eltz, 2009; Rosser *et al.*, 2019; Byers *et al.*, 2021). In some species systems, closely related allopatric or parapatric species even produce relatively similar pheromone blends yet very different blends from their sympatric close relatives (McElfresh and Millar, 2001; Rosser *et al.*, 2019). Particularly interesting, yet relatively poorly investigated from an evolutionary standpoint, is the evolution of diverse pheromone blends in evolutionary radiations. Studies on this topic are scarce and of variable scope. For example, an analysis of basal moth taxa revealed that the Δ11-desaturases needed for production of many classes of moth and butterfly pheromones were recruited before the radiation of Lepidoptera, serving as a key innovation that facilitated the evolution of diverse pheromone blends in this insect order by making new types of compounds accessible (Δ11 desaturated fatty acid derivatives) (Liénard *et al.*, 2008). Of somewhat smaller scope are some studies on the orchid bee (*Euglossini*) radiation, wherein it was found that sex pheromones diversified fast yet remained species-specific, with different species evolving different compound combinations to fill up signalling niche space (Zimmermann, Ramirez and Eltz, 2009). Orchid bee chemical signatures also show signs of reproductive character displacement in sympatric sister species pairs. Both raid diversification and reproductive character displacement are markers of

traits under sexual selection involved in reproductive isolation (Symonds and Elgar, 2008; Zimmermann, Ramirez and Eltz, 2009; Weber *et al.*, 2016).

Additionally, the genetic mechanisms underlying pheromone divergence across radiations and across species pairs are poorly resolved. But it is clear that reproductive isolation can be achieved via disparate sets of genes that control different aspects of chemical ecology (Groot, Dekker and Heckel, 2016). Sometimes divergence is given by differences in regulatory sequences for pheromone biosynthetic genes, as is the case for leafroller *Planotortrix* moths where species differences are given by deactivation of a single desaturase gene (Albre, Steinwender and Newcomb, 2013). In other cases, genetic changes in receptor proteins produce different responses to different blends in the receiving individual, a phenomenon that has been documented in two sister species of *Heliothis* moths, where males respond only to their conspecific females' chemical blend (Gould *et al.*, 2010). In yet other cases, pheromone differences between closely related species map to genes directly involved in FA metabolism, a situation that has been documented in *Ostrinia* moths and *Heliconius* butterflies (Lassance *et al.*, 2010; Byers *et al.*, 2020, 2021). These different types of genes may also act in combination to achieve pheromone differentiation. Though information on their evolutionary patterns and genomic architecture may still be lacking in many areas, there is enough published material on pheromones to hypothesize that were they to be under divergent sexual selection in a scenario of speciation-with-gene-flow, they would be suitable reproductive barriers to hybridization, and therefore we might expect to find pheromone biosynthesis or perception-related loci as speciation genes, in hotspots of genomic divergence.

The *Heliconius* system

Heliconius is a neotropical genus comprising over 40 species of butterflies (many of which include multiple races) that arose from a recent evolutionary radiation beginning ~10Mya (Kozak *et al.*, 2015). They are part of the exclusively neotropical Heliconiini tribe, including their less species-rich sister genus *Eueides,* from which they diverged ~18.4Mya, and the small "basal" genera *Podotricha, Dryadula, Dryas, Philaethria, Agraulis* and *Dione* (Kozak *et al.*, 2015). The first four of these all descend from the same common ancestor, but *Agraulis/Dione* make up a sister group to *Eueides/Heliconius* (sometimes called "advanced" genera, though this terminology is often considered obsolete), from which they diverged ~23.8Mya (Kozak *et al.*, 2015). All of the Heliconiini feed on *Passiflora* plants as larvae and are able to metabolise and sequester the cyanogenic compounds produced by their host plants' leaves within their own tissues, making them unpalatable both as larvae and as adults (Brown, 1981). Most Heliconiini species are known for their aposematism: bright and strongly contrasting colour patterns are common across the whole tribe that have evolved to signal this unpalatability to

predators (Brown, 1981). This group of butterflies has been studied extensively to answer a wide range of ecological and evolutionary questions, providing important examples of ecological sympatric speciation, speciation-with-gene-flow, convergent evolution, stable polymorphisms, adaptive introgression and hybrid speciation, has made it a very popular model organism in the field of evolution and speciation (Merrill *et al.*, 2015), In this section, I will introduce some important aspects of *Heliconius* ecology and genomics before moving on to explaining what we know about their chemical signalling.

Heliconius ecology

Heliconius is a unique genus among the Heliconiini not only for its large number of species, but for two novel adaptations that are exclusive to them within the tribe, and also very seldom seen across the Lepidopteran radiation as a whole. The two major *Heliconius* innovations are pollen feeding (Gilbert, 1972) and pupal mating (Gilbert, 1976).

Pollen feeding is a behaviour that all *Heliconius* species have in common, and is considered a key innovation in the genus (Young and Montgomery, 2020). Instead of feeding exclusively on nectar and other fluid substances, these butterflies have evolved the ability to pick up large amounts of pollen and masticate it with their proboscis until it is fully digested, gaining nutrients such as amino acids in the process that are normally inaccessible to other Lepidoptera (Gilbert, 1972). This behaviour has had a strong impact on *Heliconius* life history. The extra nutrients gained from pollen feeding allowed them to evolve a long adult lifespan of six months, the longest of any Lepidopteran when not taking periods of diapause into account (Gilbert, 1972; Boggs, 1981). The long lifespan is coupled with a much larger brain compared to other butterflies and moths, and an enhanced learning ability and visual memory (Montgomery, Merrill and Ott, 2016; Toure *et al.*, 2020; Young and Montgomery, 2020) that allows them to memorize and recall information on resource availability within their territory.

Meanwhile, the other major *Heliconius* innovation, pupal mating (Gilbert, 1976), evolved as *Heliconius* split into two lineages ~12Mya, and became exclusive to one of the two, including the *H. sara-H. sapho* and *H. erato* subclades and representing nearly half of all *Heliconius* species (Kozak *et al.*, 2015). Adult males of pupal-mating species are attracted by pheromones emitted by female pupae of their own species (Estrada *et al.*, 2010). Once they locate a female pupa, males perch and guard them against any competing males or other approaching butterflies, and mate with them as soon as emergence commences (Gilbert, 1976). This behaviour is not unique across all Lepidoptera, but is not observed at all in the other *Heliconius* lineage, where sexually mature males engage in courtship with sexually mature females (Klein and de Araújo, 2010; Vanjari *et al.*, 2015; Darragh *et al.*, 2017; Mann *et al.*, 2017) with the consequence being that females of the “free-mating” clade can occasionally re-mate

more often than those in the pupal-mating clade (Estrada *et al.*, 2011; Walters *et al.*, 2012). However, even within the pupal-mating clade, it is a facultative behaviour dictated by availability of female pupae (Thurman *et al.*, 2018).

The most striking aspect of *Heliconius* ecology is their warning patterns, which first attracted the attention of tropical researchers to this genus (Bates, 1862). All Heliconiini are unpalatable and brightly coloured, but to complement this, *Heliconius* partakes in Müllerian mimicry, as distantly related species convergently evolve the same colour patterns to share the cost of educating predators (Bates, 1862; Mallet and Gilbert, 1995). Perhaps surprisingly, often *Heliconius* species belong to different mimicry rings across their distributions in the Amazon and in Central America (Mallet and Gilbert, 1995). This applies especially to species with large distributions like *H. melpomene* and its distantly related mimic, *H. erato* (Mallet and Gilbert, 1995; Mallet and Joron, 1999), or *H. numata* which is also locally polymorphic (Joron *et al.*, 1999). Hybridization between different colour pattern is disadvantageous, as predators do not recognize the disrupted patterns seen in hybrids (Naisbit, Jiggins and Mallet, 2001).

Genomics of speciation in *Heliconius*

Most Heliconiini sister species pairs are parapatric or sympatric, with 32-95% range overlap, suggesting that speciation-with-gene-flow may be widespread across the tribe (Rosser *et al.*, 2015) and making these butterflies good candidates for the study of speciation genes. Research on *Heliconius* genetic divergence is facilitated by the availability of high-quality reference genomes for *H. melpomene* and *H. erato* (Dasmahapatra *et al.*, 2012; Davey *et al.*, 2016; Van Belleghem *et al.*, 2017), making the analysis of whole genome sequence data and approaches such as GWAS and QTL mapping easier. Low levels of genomic divergence, a clear sign of gene flow between taxa, have been detected across the genome in between several *Heliconius* taxa, and in all of these cases, it pointed to greater admixture between sympatric pairs: among races of *H. melpomene* (Nadeau *et al.*, 2012), races of *H. erato* (Martin and Van Belleghem, 2017), across *H. melpomene, H. cydno* and *H. timareta* (Martin *et al.*, 2013) and between *H. elevatus* and *H. pardalinus* (Dasmahapatra *et al.*, 2012). However, levels of gene flow were not uniform across the genome, revealing major divergence hotspots (Dasmahapatra *et al.*, 2012; Nadeau *et al.*, 2012; Martin *et al.*, 2013). These represent the areas under divergent selection, where one might expect to find "speciation genes"; indeed, colour pattern genes are found within these islands of divergence in the face of gene flow (Dasmahapatra *et al.*, 2012; Nadeau *et al.*, 2012; Martin *et al.*, 2013). The colour divergence islands are mainly located on chromosome 10, 15 and 18 and contain well known patterning genes such as *optix* for the arrangement of red/orange colour pattern elements (Reed *et al.*, 2011), *cortex* for yellow/white elements (Nadeau *et al.*, 2016)

and *WntA* that contains several loci encoding positional information for various pattern elements (Martin *et al.*, 2012).

In some species, the locus controlling mate choice in *Heliconius* is known to be genetically linked to colour pattern loci (Jiggins *et al.*, 2001; Kronforst *et al.*, 2006; Merrill *et al.*, 2011), so female mate preference is tightly connected to male colour pattern; indeed, female *Heliconius* show a strong preference for individuals with their own colour pattern (assortative mating) in some pairs, including *H. melpomene/H. cydno, H. cydno/H. pachinus, H. heurippa* races, and *H. elevatus/H. pardalinus* (Jiggins *et al.*, 2001; Kronforst *et al.*, 2006; Mavárez *et al.*, 2006; Rosser *et al.*, 2019). This makes colours the most widely studied reproductive barrier in the genus. However, colour pattern alone does not fully explain reproductive isolation (Jiggins *et al.*, 2001; Jiggins, 2008; Mérot *et al.*, 2015), as co-mimetic sister species do not interbreed either (Mérot *et al.*, 2015). This means that some loci other than colour pattern ones must be under divergent sexual selection, an argument made all the more convincing by the fact that there are *Heliconius* sister pairs that diverge in aspects of their ecology unrelated to the colour pattern, such as host plant, pheromone composition or flight dynamics (Mérot *et al.*, 2013, 2015, 2017; Rosser *et al.*, 2015, 2019; Queste, 2020).

What other aspects of *Heliconius* divergence can be mapped to divergence hotspots between sympatric sister pairs? Naturally, we need to turn our attention to other traits that show strong divergence between sympatric species pairs. Pheromones have been proposed before as a possible reproductive barrier (Merrill *et al.*, 2015), and are known to be strongly differentiated between at least some sister species pairs (Mann *et al.*, 2017; Rosser *et al.*, 2019; Byers *et al.*, 2020), but have received relatively little attention in research.

Sex pheromones in *Heliconius*: what do we know?

Compared to colour patterns, *Heliconius* pheromones are poorly characterised and understood. This is probably due to our own sensory bias: colour patterns can be visually analyzed, and their characterization was already possible at the time of Bates' initial observations on the genus 160 years ago (Bates, 1862). Chemical emissions on the other hand require dedicated instrumentation; gas chromatography-mass spectrometry (GC-MS) instruments, used to analyze and identify volatile components of a mixture, have much improved in the past decades, making analysis of animal scents easier (Sparkman, Penton and Kitson, 2011), and opening the door to research on *Heliconius* chemical ecology. The first observation on *Heliconius* pheromones dates back to 1976 when mated *H. erato* females were observed to emit a pungent male-contributed odour with apparent repellent effect on other males (Gilbert, 1976). Experimental research on *Heliconius* pheromones however only truly began in the last two decades, as the first publication on the topic dates only to 2007 (Schulz *et al.*,

2007). A search on webofscience.com for "*Heliconius* pheromones" shows 25 publications on the topic from 2008 to 2021, a number that has been picking up in recent years but still far removed from the hundreds of publications on *Heliconius* colour patterns. These publications form the foundation for this thesis, and are all extremely valuable in this early phase of *Heliconius* pheromone research. Aside from one paper on female pupal scent emissions (Estrada *et al.*, 2010), all of them have been about male adult pheromones, specifically two kinds of them: antiaphrodisiacs and male sex pheromones (MSPs), also known as courtship pheromones.

Antiaphrodisiacs. The first papers on *Heliconius* chemical ecology concern their abdominal secretions (Schulz *et al.*, 2007, 2008; Estrada *et al.*, 2011). Schulz et al. (2007) identified the components of abdominal emissions from *H. cydno* and *H. pachinus,* consisting of complex fatty acid (FA) derivatives known as macrolides (Schulz *et al.*, 2007). A year later, it was discovered that in *H. melpomene* these abdominal scents can have an antiaphrodisiac function (Schulz *et al.*, 2008). The complex FA-derived compounds were found in this species too, albeit not the same kinds observed in *H. cydno* and *H. pachinus,* and form a relatively heavy (not very volatile), complesx "matrix" with mild antiaphrodisiac properties, but are accompanied by one extremely abundant and very volatile compound, a terpene called (E)-β-ocimene, with potent behavioural effects (Schulz *et al.*, 2008). This compound, accompanied by the heavier matrix, is delivered from male butterflies to female butterflies during mating and discourages further mating events by signalling that the female is already mated (Schulz *et al.*, 2008; Estrada *et al.*, 2011), and it was indeed the same antiaphrodisiac compound first observed by Gilbert in *H. erato* (Gilbert, 1976). (E)-β-Ocimene is relatively widespread as a behaviourally active compound among insects (Honda, 1981; Chaturvedi, Singh and Sane, 1990; Keegans *et al.*, 1993; Sun *et al.*, 2010; Eller and Palmquist, 2014), but it is also extremely common as a floral scent (Knudsen *et al.*, 2006). It would be intuitive, then, to classify (E)- β-ocimene as a host plant-derived compound, but instead, *Heliconius* produces it *de novo* using its own terpene synthase (Darragh, Orteu, *et al.*, 2019).

The antiaphrodisiac composition observed in *H. melpomene* is a good model for the genus in general. Genital blends in *Heliconius* are commonly comprised of one or a few early-eluting volatiles accompanied by a heavier matrix (Estrada *et al.*, 2011). Despite this common arrangement, the composition of the genital blend is wildly variable across species (Estrada *et al.*, 2011). However, despite showing the explosive diversification patterns observed in some reproductive barrier traits, antiaphrodisiacs are not good candidates for a role in prezygotic reproductive isolation, as they are only effective after mating. Instead, it is believed that their great variability in composition is due to sexual conflict, as antiaphrodisiacs are beneficial to males but not to females (Estrada *et al.*, 2011), so their composition is under weaker selection. It has also been shown, on a small subset of free-mating and pupal-mating species, that the former show greater variation in their antiaphrodisiac

composition, likely driven by greater male-male competition among free-mating individuals compared to the pupal mating clade where female choice is less relevant (Estrada *et al.*, 2011). This works based on the assumption that polyandry, and so male-male competition, is more commonplace in free-mating species, whereas pupal-mating species are largely monoandrous. However there is little evidence that the difference in mating strategy between the two *Heliconius* lineages is truly strong enough to explain a shift in evolutionary rate (Walters *et al.*, 2012; Thurman *et al.*, 2018).

While the function of the early eluting compound and matrix components has been behaviourally tested in *H. melpomene* (and to a lesser extent in *H. erato*), meaning their antiaphrodisiac effect has already been experimentally verified (Schulz *et al.*, 2008; Klein and de Araújo, 2010), the same does not apply to the great diversity of other compounds observed in other species, whose antiaphrodisiac function is only putative at the time of writing. Therefore, they shall be more cautiously referred to as "genital compounds" or "genital scents".

Male sex pheromones (MSPs). While the earlier experimental studies on *Heliconius* pheromones focused on their genital extracts, male sex pheromones (MSPs) have since received much attention. These scent bouquets are emitted by males during courtship via specialized hindwing scales called androconia (Brown, 1981; Darragh *et al.*, 2017; Mann *et al.*, 2017), actually consisting of living secretory cells that store fatty substances (Darragh *et al.*, 2017). Wing colour pattern is not the only factor female *Heliconius* use to assess whether to mate with a male (Jiggins, 2008), and blocking the androconia from secreting MSPs negatively affects the male's mating success, meaning that even when presented with a conspecific colour pattern, the female is deriving some important information from the pheromone blend (Darragh *et al.*, 2017), though female criteria for mate choice based on male scent are still entirely unexplored. The information contained in MSPs is surely enough for species recognition, and there are even cases of co-mimetic species relying on their pheromone composition for mate choice (González-Rojas *et al.*, 2020).

Heliconius MSPs are not as complex as their antiaphrodisiac blends, comprising generally fewer compounds of fewer classes, but they are still very variable across the genus, while being species-specific or even race-specific (though chemical profiles tend to still be similar between races of the same species) (Mann *et al.*, 2017; Darragh *et al.*, 2020). The androconial blends of several species and races have been characterized, and it was found that, as is common among butterflies, they are comprised of fatty acid derivatives, including blends of alkanes, alkenes, esters, alcohols and aldehydes, both saturated and desaturated, accompanied by plant derivatives, most commonly phytol and its degradation product (Schulz, Yildizhan and van Loon, 2011) and lignin-derived aromatic compounds such as homovanillylalcohol and syringaldehyde (Mann *et al.*, 2017). The fatty acid

derivatives are all produced via the same pathway observed across other Lepidoptera (Reed *et al.*, 1994, 1995; Tillman *et al.*, 1999; Liénard *et al.*, 2014; He *et al.*, 2017; Mann *et al.*, 2017), starting from palmitic acid as the primary precursor, which can then be elongated into stearic acid, at which point a double bond may be introduced by a Δ9-desaturase, leading to oleic acid. Both stearic and oleic acid can further be elongated and then transformed into the full range of fatty acid derivatives, respectively saturated or desaturated, and no final product is mutually exclusive of the others, meaning any *Heliconius* species may produce scent bouquets containing any combination of saturated or desaturated fatty acid products, be they alkenes, alkanes, alcohols, aldehydes or esters (Mann *et al.*, 2017). Many of these reactions are mediated by enzymes with low substrate specificity (Reed *et al.*, 1994, 1995; Tillman *et al.*, 1999; Mann *et al.*, 2017), and at least in *H. melpomene-H. cydno,* utilizing F1 hybrids and backcrosses, some of the putatively involved enzymes have been genomically mapped to QTLs located on chromosome 20 (Byers *et al.*, 2020, 2021).

This all has potentially big implications for chemical signal-based reproductive isolation in *Heliconius,* as the inherent flexibility of this pathway and the wide range of different products it can lead to may mean that even closely related *Heliconius* species may be able to evolve divergent pheromone blends in relatively short evolutionary times. This further invigorates interest in MSPs as reproductive barriers in *Heliconius.*

Aims of this study

Chemical signals have the potential to be important drivers of speciation in *Heliconius* as in many other taxa. In this study, I use these butterflies as a model system to investigate various aspects of pheromone evolution, how it may impact speciation and on the other hand, how it might be shaped by ecological factors. I am looking to not only provide evidence for the role of MSPs as reproductive barriers, but also to provide a foundation for future *Heliconius* pheromone-related studies.

First, I perform a phylogenetic analysis on MSP composition not only in *Heliconius,* but in the entire Heliconiini tribe, in order to gain a broad perspective of their diversification and macroevolutionary patterns. This is the most inclusive analysis of Heliconiini pheromones to date, and I characterize the previously undocumented chemical extracts of several species in the process. In this first part, I focus on questions pertaining to MSPs' putative role as reproductive barriers across the *Heliconius* radiation, informed by past studies on reproductive isolation in evolutionary radiation: are MSPs able to diversify quickly, yet remain species-specific? Are there any differences in pheromone profiles or biosynthesis between "basal" genera and *Heliconius* that may have affected the MSP diversification rates? And most importantly, are androconial blends more strongly differentiated in sympatric sister pairs? (Chapter 1).

I then focus on one sympatric sister pair in particular, *H. elevatus* and *H. pardalinus,* to perform a quantitative trait locus analysis. To perform this analysis, I take advantage of the species' ability to produce fertile hybrids and use F2 crosses to map their pheromone differences to the genome and test whether their pheromone loci fall within hotspots of genomic divergence. Having identified confidence intervals for some of the main compounds that describe species differences, I then look for genes related to the fatty acid metabolic pathway in the regions of interest (Chapter 2).

Having performed a comprehensive analysis of androconial extracts, I then move on to characterizing genital extracts across the Heliconiini tribe, providing a useful comparison between MSPs, which are good candidates as drivers of reproductive isolation, and antiaphrodisiacs, which are likely to be uninvolved. Using new multivariate phylogenetic analysis techniques, I further investigate whether mating strategy (free mating vs. pupal mating) has affected genital scent diversification, and in the process, I also attempt to test the putative role of volatile compounds as behaviourally active compounds across the tribe, theorizing that the potent antiaphrodisiac effect of (E)-β-ocimene may also be present in other abundant volatiles (Chapter 3).

Finally, I perform an analysis on the effect of age on pheromone composition in two species representative of the two *Heliconius* lineages: *H. atthis* (free mating) and *H. charithonia* (pupal mating), in order to understand some of the within-species variation observed in pheromone profiles, and to assess whether the timing of pheromone production in these species depends on their mating strategy (Chapter 4). Age is only one of many basic variables whose effect on chemical emissions have never been investigated before, that may affect the butterflies' chemical signalling and alter their signatures to an extent that females may be able to perceive, providing them with information that they may use for mate choice. This could provide the foundation for future studies on the broadly unexplored topic of female choice criteria in *Heliconius.*

Chapter 1: Activation of an ancestral pheromone biosynthetic pathway contributes to diversification in *Heliconius* butterflies

Abstract

During courtship, male butterflies produce androconial secretions containing sex pheromones that affect female choice by providing information about species identity and possibly fitness. Due to their function in species recognition, sex pheromones may act as reproductive barriers contributing to speciation. However, in general, the role of pheromones in speciation remains poorly studied. Although neotropical *Heliconius* butterflies are a model system in the study of speciation, and there are a few studies on their pheromones , the effect of male sex pheromones at a macroevolutionary level has not been investigated. We use gas chromatography/mass spectrometry to characterise male androconial secretions in 33 of the 69 species in the Heliconiini tribe. These pheromone-containing blends were found to be species specific, consistent with a role in reproductive isolation. We detect a significant increase in the rate of pheromone blend diversification at the root of the most speciose genus, *Heliconius*; a consequence of *Heliconius* and *Eueides* species using the fatty acid metabolic pathway to unlock more complex chemical blends compared to species in the basal Heliconiini genera, whose pheromone blends are dominated by plant derivatives. A comparison of the pheromones of 10 sister-species pairs demonstrates a striking pattern of greater disparity in pheromone blends with increasing geographic overlap between the species pairs; a pattern consistent with a scenario of character displacement or reinforcement between sympatric compared to allopatric taxa. Taken together, these results demonstrate for the first time that diversification of male sex pheromones are important for initiating and/or maintaining species reproductive barriers across this group of butterflies, providing an example of how re-activation of an ancestral trait, the co-option of the fatty acid metabolic pathway for pheromone production, can facilitate rapid speciation.

Introduction

Invasion of a new ecological niche can lead to evolutionary radiation, leading to enhanced ecological or morphological disparity among the resulting species as they diverge to exploit different aspects of the niche (Yoder et al., 2010). This increased disparity can result from novel and unique evolutionary changes, often called key innovations, that accompany high diversification rate in the clade in which they are found (Miller, 1949; Heard and Hauser, 1995; Rabosky, 2017). These innovations are thought to facilitate diversification by unlocking "ecological opportunity": the ability for species to diverge into underutilized ecological niches (Hunter, 1998; Yoder et al., 2010; Dumont et al., 2012). Niche expansion and the resulting ecological opportunity can make room for new adaptations and in turn pave the way to an evolutionary radiation in a number of different ways (Yoder et al., 2010; Dumont et al., 2012). In theory, proving that trait innovations are linked to diversification involves showing that the clade bearing the innovation has undergone faster diversification than its relatives, that such diversification has occurred in tandem with the evolution of a trait, and that the trait has also triggered adaptive phenotypical changes in the organism (Dumont, 2011). In practice, demonstrating that ecological opportunity granted by innovations has led to increased diversification has yielded mixed results depending on the organism and adaptation in question (Hunter and Jernvall, 1995; Hagen and Kadereit, 2003; Alfaro et al., 2009; Dumont et al., 2012; Silvestro et al., 2014). The difficulty is often with correctly identifying a trait as a key innovation, since an innovation such as a change in diet (Lovette, Bermingham and Ricklefs, 2002; Konow *et al.*, 2017; Young and Montgomery, 2020) is in reality a composite of multiple traits that have evolved simultaneously.

However, a novel trait does not need to be part of a key innovation to open up new niches at a finer level. The recovery of a previously lost trait, while by definition not an innovation, may also have a similar effect, and testing its effect on clade diversification may pose fewer issues than larger changes that arise with the modification of multiple traits. Niches are often thought of in an ecological sense, as the range of different positions occupied by organisms within the ecosystem, such as the variety of food sources, feeding substrates, or habitat preferences (Polechová and Storch, 2019). But with traits involved in reproductive isolation, niches can reflect the availability of signal space. Reproductive character displacement may lead to the evolution of different mating signals that do not overlap with those of closely-related species (Brown and Wilson, 1956; Blair, 1974; Pfennig and Pfennig, 2009). The rapid acquisition of such non-overlapping signals is facilitated when the traits in question can evolve in a relatively unconstrained manner.

Chemical odours are important intra and interspecific signals across the tree of life, conveying information about fitness, reproductive status, species and individual identity, as well as serving as alarm or aggregation calls, among a variety of other purposes (Regnier and Law, 1968). Pheromones are a widespread form of chemical communication that elicit behavioural responses in receiving individuals (Karlson and Lüscher, 1959; Symonds and Elgar, 2008; Ginzel, 2009; Yew and Chung, 2015). They are extremely variable among species, both in function and in composition, and are an essential communication system in unicellular and multicellular organisms (Fields, 1990; Gassmann et al., 1995; Jaenicke and Starr, 1996; Wirth et al., 1996; Dulac and Torello, 2003; Williams et al., 2007; Symonds and Elgar, 2008; Frenkel et al., 2014; Luporini et al. 2015). Most macroevolutionary studies of signalling tend to focus on more easily quantifiable signalling, like traits involved in morphology (e.g. Freeman, 2000; Jousselin et al., 2004; Mahler et al., 2010) or visual and auditory communication, such as coloration and song in birds (Maia et al., 2013; Mason et al., 2017; Matysioková et al., 2017; Pearse et al., 2018; Merwin et al., 2020; Beltrán et al., 2021). Therefore, despite the ubiquity of pheromone signalling, the difficulty of testing and quantifying such signals and their behavioural effects means that pheromone macroevolution is a largely under-explored topic, though studies on the topic do exist, both in insects (Symonds and Elgar, 2004; Symonds and Elgar, 2008; Symonds and Gitau-Clarke, 2016; Weber et al., 2016) and in other groups (Buchinger et al., 2017; Dong et al., 2020). While the role of pheromones in reproductive isolation has been explored before (Smadja and Butlin, 2009; Shahandeh et al., 2018) there is a general lack of studies on the expansion and exploitation of odour niche space to facilitate reproductive isolation.

Neotropical *Heliconius* butterflies are representative of the bias towards visual traits. Over a century of research (Bates, 1862) on this genus (Baxter *et al.*, 2008; Dasmahapatra *et al.*, 2012; Merrill *et al.*, 2015) has provided answers to many evolutionary questions pertaining to the role that the vivid, aposematic wing patterns that characterize the radiation play in adaptation and speciation (Jiggins *et al.*, 2001; Salazar *et al.*, 2010; Dasmahapatra *et al.*, 2012; Mallet and Dasmahapatra, 2012; Martin *et al.*, 2013; Jiggins and Lamas, 2017; Mérot *et al.*, 2017), as well as adaptive introgression (Jiggins *et al.*, 2008; Salazar *et al.*, 2010; Dasmahapatra *et al.*, 2012; Pardo-Diaz *et al.*, 2012), and the role of "magic traits" (Merrill *et al.*, 2012; Mérot *et al.*, 2015) in animal evolution. By contrast, relatively little attention has been paid to other traits involved in reproductive isolation, including the role of chemical signalling (Merrill *et al.*, 2015). Yet reproductive isolation among *Heliconius* species cannot be fully explained by colour pattern alone, and other ecological traits, including male sex pheromones, also likely have key roles (Jiggins, 2008; Mérot *et al.*, 2015; Queste, 2015; Rosser *et al.*, 2019). Nonetheless, from the early observations of Bates (1862) on the possible role of mimicry on the increased rate of speciation in *Heliconius*, it took until 2008 for the first *Heliconius* pheromone to be identified (Schulz

et al., 2008) and until 2011 for the first paper on the evolutionary dynamics of some of their chemical signals to be published (Estrada et al., 2011).

Male sex pheromones are good candidates as reproductive barriers in butterflies due to their effect on mate choice. In *Heliconius*, these are produced by structures known as androconia located on the male hindwing (Supp Fig. 2) (Darragh et al., 2017; Mann et al., 2017), and the chemicals blends are mostly comprised of a mixture of plant and fatty acid (FA) derivatives (Mann et al., 2017). The plant derivatives are acquired and/or transformed from plant products, and the fatty acid derivatives are produced endogenously from ubiquitous fatty acid precursors stored in the butterfly body (Mann et al., 2017; Darragh, Byers, et al., 2019). Both types of compounds are commonly found in insect pheromones (Ando et al., 2004). The importance of these androconial compounds in mate choice is demonstrated by the fact that male *Heliconius* with blocked androconia are largely unsuccessful at mating compared to the untreated males (Darragh et al., 2017), and at least one compound has been shown to have a function in mating through both behavioural and electrophysiological assays in *H. melpomene* (Byers et al., 2020). Furthermore, in the co-mimetic sister species *H. melpomene* and *H. timareta,* females prefer males perfumed with conspecific rather than heterospecific androconia secretions (Mérot et al., 2015), demonstrating the importance of scent in mating.

Rapid evolution of divergent pheromone blends may therefore be an important driver of reproductive isolation. The fatty acid metabolic pathway, responsible for much of the androconia pheromone blend in *Heliconius* (and the pheromones of other Lepidoptera) (Liénard et al., 2014), is modular in nature, comprising many reactions that can be activated or deactivated to produce different cocktails of final products (Mann et al., 2017) (Supp. Fig. 1.1). This pathway thus has the potential to be evolutionarily labile, allowing species to rapidly occupy non-overlapping niches in odour signalling space. Use of the fatty acid metabolic pathway may therefore have been instrumental in shaping the role of pheromones as reproductive barriers in this group of butterflies. Similar traits involved simultaneously in sexual signalling and mating isolation, such as courtship songs in *Drosophila,* have been shown to be particularly labile and able to diversify faster than traits not involved in mating (Gleason and Ritchie, 1998).

Here we evaluate whether male sex pheromones in the Heliconiini tribe of butterflies display macroevolutionary patterns consistent with those expected from a trait involved in reproductive isolation. We first determine the degree that male sex pheromone blends are species-specific (Ord and Stamps, 2009). We then assess the rate at which pheromone blends evolve across this group of butterflies to determine whether species-rich clades have more rapidly evolving pheromone blends. Finally, we test the hypothesize that if pheromone blends are playing a key role in reproductive isolation then pheromone divergence between sister species should be correlated with the extent of

geographic range overlap between them.

Materials and Methods

Gas chromatography/mass spectrometry (GC/MS) analysis of androconial pheromones

We analyzed the wing extracts of 179 male and 66 female adult butterflies from 33 of the 69 Heliconiini species, obtained from either wild or captive-bred populations (Figure 1.1, Supp. Table 1.2). An average of 6 samples were obtained per species (minimum 1; maximum 15). Wild butterflies were sacrificed following capture, while captive-bred butterflies were sampled on reaching sexual maturity (at least 10 days after eclosion). Two wing tissues were dissected from males: the androconia and a control region of the wing not involved in pheromone secretion (Supp. Fig. 1.2). The equivalent wing regions were sampled in females when they were available. Following dissection, the tissues were extracted using 200µl of dichloromethane with a 2-acetoxytetradecane (1 ng/µl) internal standard.

Tissue extracts were analyzed on a GC-MS system which consisted of a 7890A GC-System coupled with an MSD 5975C Mass Analyzer (Agilent Technologies, Santa Clara, USA) fitted with a HP-5MS column (50m length 0.25 mm I.D., 0.25 µm film thickness; Agilent Technologies). The ionization -method was electron ionization with an electron energy of 70 eV. The instrument conditions were as follows: inlet pressure 9.79 psi, He 20 mL min^{-1}, injection volume 1 µL. For most samples, the GC was programmed with the following temperature ramp: start at 50°C, increased by 5°C min^{-1} to a maximum temperature of 320°C, then hold for 5 min (total run time 69 min). Most samples collected in Panama were analyzed in the Chemical Ecology laboratories at the STRI Earl S. Tupper Center in Panama City in 2018. *H. timareta* and *H. eleuchia* samples were run by Kathy Darragh (University of Cambridge), with the same method and on the same GC-MS system, between late 2017 and early 2018. A subset of samples was run with using the same conditions and on the same instrument model at Braunschweig Technische Universität in Braunschweig (Germany), although some of these were analyzed using a slightly different temperature ramp, starting at 50°C, then increasing at 10°C min^{-1} to a maximum temperature of 320°C (total run time 45 min).

The program AMDIS (Mallard and Reed, 1997) was used to quantify and identify all compounds with gas chromatographic retention indices (RI) between 1031 (4- hydroxypentan-2-one) and 2900 (nonacosane) through comparison of mass spectra and RI values with synthetic samples and mass spectrometric databases (see Mann et al. 2017 for full details of the GC/MS protocol and analysis). Compound concentrations obtained from this analysis were transformed to logged percentages (of the total amount of androconial contents) for use in most subsequent analyseanalyses.

We used this data to generate three datasets:

Dataset A: An unfiltered **dataset C**omprising all compounds detected in each individual via GC/MS, with the exception of known contaminants and structural components (RI≥2600).

Dataset B: A filtered subset of **Dataset A** comprising all compounds that were present in significantly different amounts in the androconia compared to the male control tissue in at least one species, as determined using one-tailed paired t-tests.

Dataset C: Derived from **Dataset B**, this dataset retains for each species only the compounds present in significantly different amounts between androconia and male control tissues. We term these compounds putative male sex pheromones (MSPs). For *H. burneyi,* where only one male was sampled, we retained all compounds that only appeared in the androconia but not the control.

Most of the subsequent tests were run on both unfiltered and filtered datasets.

Pheromone diversity and species identity

For each species, using logged percentages of compounds for both the filtered dataset (**dataset B**) and putative MSPs (**dataset C**), we calculated the richness, evenness and diversity (Shannon Index) of androconial contents using vegan (Dixon, 2003). These statistics were compared between the traditionally named basal and "advanced" genera (*Heliconius* + *Eueides*) using Welch's 2-sample t-tests. Differences in putative MSP (**dataset C**) composition between species were visualised using non-metric multidimensional scaling (NMDS) on logged percentages. The NMDS was calculated on Euclidean distances using the metaMDS function in vegan (Dixon, 2003), based on 1000 random starts and 5 dimensions. The number of dimensions (k) was chosen based on the stress level; 5 axes give an accurate representation of the data without overly inflating the number of dimensions.

To test whether pheromone blend is consistent within species as expected of a trait involved in reproductive isolation, we assessed the species-specificity of pheromone composition. Concentrations from the filtered dataset (**dataset B**) were converted into logged percentages, and pairwise Euclidean distances used as a measure of dissimilarity between individuals. Species-specificity was calculated using the ADONIS function (permutational analysis of variance based on distance matrices) from the vegan package (Dixon, 2003), and significance was determined via 10000 free permutations.

As a first step towards assessing the consistency of pheromone blend among related species, we also measured the clade-specificity of pheromone composition. The clades used are somewhat arbitrary but widely accepted (Kozak *et al.*, 2015), and are depicted in Figure 1. Concentrations from the filtered dataset (**dataset B**) were averaged by species and converted into logged percentages. The Euclidean distance between species' average was used as a quantifier of dissimilarity. Clade-specificity and significance were calculated as above.

Diversification rate of pheromone blends

We used the published ultrametric Bayesian tree for the Heliconiini (Kozak et al. 2015) in all phylogenetic analyseanalyses. In species where samples from multiple races were available, a single race (*H. erato notabilis, H. melpomene plesseni,* and *H. pardalinus butleri*) was chosen as representative of the species due to either good availability of samples or known branch length. This was to avoid the effect of polytomies, which can compromise the results of phylogenetic tests (Walsh *et al.*, 1999). We used the filtered dataset (**dataset B**) for all the following phylogenetic analyseanalyses. The phylogenetic signal was quantified using the Kmult parameter (Adams, 2014) calculated with the physignal function in geomorph (Adams and Otárola-Castillo, 2013). Significance was tested via 10000 random permutations of the data between the tips of the phylogeny. To assess whether specific pathways or families of compounds were more phylogenetically conserved than others, Kmult was also calculated for the following subsets of compounds: fatty acid derivatives, plant derivatives, and the lignin and phytol families of plant derivatives separately. The former two were chosen due to their different biosynthetic origins, whereas the latter two were chosen as the two largest and most widespread groups of plant derivatives in the dataset.

Using the R package mvMORPH (Clavel et al., 2015), alternative evolutionary models were fitted to test different tempos of evolution, as well as shifts in the rate of pheromone evolution across the Heliconiini, and to identify any detectable bursts in pheromone diversification together with the nodes in the phylogeny they were associated with. The data used for this consisted of five-dimensional NMDS scores averaged within each species and based on **dataset B**. The models tested include three non-shift models (Brownian motion, early burst and Ornstein-Uhlenbeck) as well as models with rate or mode shifts at two different points in the phylogeny: *Eueides* + *Heliconius*, and *Heliconius* (Table 1.1). The reasons for these choices were as follows: the *Eueides/Heliconius* split is the base of the most diverse genera, and corresponds to the introduction of the FA metabolic pathway for pheromone production as found in our results from this study; the base of *Heliconius* marks the beginning of the *Heliconius* evolutionary radiation. The models were compared using the corrected Akaike Information Criterion (AICc) and the relative Akaike weight. The compare.evol.rates function from geomorph was then used on **dataset B** as well as on its 5 NMDS axes to assess the shift in rate for the best fitting model, based on 10000 random permutations.

Shifts tested with mvMORPH were chosen based on *a priori* information on the clade's evolutionary history, but at the same time the R package MOTMOT (Thomas and Freckleton, 2012) was used to find rate shifts in an unguided manner using the function transformPhylo.ML with the "TraitMedusa2" ("tm2") model. To avoid the strong effect potentially posed by any strongly differentiated sister species pairs, which would produce a bias for evolutionary shifts at the tips, the minimum clade size

for inferred rate shifts was set as 3. The calcCutOff function was used to calculate an appropriate AICc cutoff for the tm2 model, based on 1000 simulations with the minimum clade size threshold taken into account. This AICc cutoff was used to determine the best fitting TraitMedusa2 output between the basic no-shift Brownian motion model and evolutionary rate shift models.

Comparison of sister species pairs

If pheromones act as reproductive barriers, sympatric sister species pairs are expected to produce more strongly divergent pheromone blends as a result of character displacement, compared to sister pairs with little or no range overlap. We used a GLM with Gaussian distribution to test the effect of range overlap as reported in (Rosser *et al.*, 2015) and divergence times (branch length from (Kozak *et al.*, 2015)) on pheromone dissimilarity (Euclidean distance, calculated using the unfiltered **dataset B** with the R package Geiger (Harmon et al., 2008)) between 10 of the 22 Heliconiini sister species pairs for which we had data (Rosser *et al.*, 2015). Divergence times between sister pairs were included to control for evolutionary time.

Correlations between compounds

The co-occurrence of compounds across multiple species may arise simply as a consequence of compounds sharing the same biosynthesis pathway. Alternatively, such a correlation may be biologically more significant, such that two or more compounds act together to elicit a behavioural response. This was previously observed to an extent in *Heliconius* antiaphrodisiac pheromones, where early eluting volatile compounds and heavier late eluting FA esters had a combined behavioural effect that was stronger than isolated early volatiles and esters (Schulz et al., 2008). Using species averages for each compound, we tested for correlations between all pairs of compounds in the filtered dataset (**dataset B**). To correct for the phylogenetic relatedness, we computed the phylogenetic variance-covariance matrix and then converted it into a correlation matrix, which was then visualized using R(corrplot) (Wei and Simko, 2017). We calculated the p-values separately as described in Revell (2017). Due to the large number of correlation tests resulting from this analysis, we utilized the Benjamini-Hochberg procedure to correct p-values and control for the false discovery rate (Benjamini and Hochberg, 1995) using the p.adjust function from the R base stats package.

Results

GC/MS analyseanalyses detected 127 compounds from 231 samples (170 males and 61 females) representing 36 Heliconiini species, 3 of which (*H. erato, H. melpomene* and *H. pardalinus*) included 2 races. From these, we removed cuticular hydrocarbons (RI≥2600) as they are ubiquitous in insect tissues and too heavy for longer range signalling. These were all long chain alkanes, and there were

only five instances where a cuticular hydrocarbon was significantly higher in proportion in the androconia compared to the control: hexacosane in *D. juno,* octacosane in *D. phaetusa* and *H. pachinus,* and nonacosane in *E. tales.* Due to the ubiquity of cuticular hydrocarbons, these differences are likely due to different amounts of sampled tissue. This left us with 117 compounds, making up **Dataset A** (Supp. Table 1.5). Three of the featured species, *H. eleuchia, H. melpomene rosina* and *H. timareta,* lacked control samples and were thus not included in any filtering stage of the analysis, leaving a total of 33 species. Only 86 of the 117 compounds were significantly higher in proportion in the androconia compared to the male control in at least one species, as per Welch's 2-sample t-test. These 86 compounds were used in the filtered analysis, making up **Dataset B**. The same 86 compounds make up **dataset C**. For **Dataset C**, however, we only retained, for each individual, compounds present in significantly higher amounts in the androconia of that species compared to the negative control, while nullifying the concentrations of other compounds that do not show that significant difference. We refer to the retained compounds for each species as "putative male sex pheromones (MSPs)", as the behavioural activity of most of them has not been ascertained. In most species, male androconia samples contained significantly more compounds, and in significantly larger quantities than both male hind wing and female controls (Supp. Table 1.3), demonstrating the role of androconia as male tissues specialised for secreting complex chemical blends (Rosser *et al.*, 2019) This is consistent with behavioural studies in several *Heliconius* species which have shown that compared to unmanipulated males, males whose androconial output has been blocked have strongly reduced mating success (Darragh *et al.*, 2017). For a list of all 117 compounds and information on which were retained in **dataset B** and C, see Supp. Table 1.5.

Qualitative comparison of advanced and basal Heliconiini

The androconia extracts of the advanced Heliconiini (*Heliconius* + *Eueides*) contained significantly more compounds than those of the basal Heliconiini (*Philaethria*/*Dione*/*Dryadula*/*Dryas*/*Agraulis*); on average 20±7.6(SD) vs. 11±3.1(SD) compounds respectively (Welch's two-sample t-test, t=4.36, df=9, p=0.002) when considering **Dataset A.** In **Dataset B** and **C**, this pattern remains unchanged, with 17±6.3(SD) vs 10±2.9(SD) compounds respectively (t=4.13, df=8, p=0.004) for **B** and 7±4.2(SD) vs. 3±1.9(SD) compounds respectively (t=3.09, df=8, p=0.01) for **C**. The evenness of compounds was not significantly different between the advanced and basal Heliconiini regardless of the filtering level (**Dataset B**: t=-1.30, df=4, p=0.27; **Dataset A**: t=-2.67, df=4, p=0.08). Likewise, the diversity (Shannon index) of compounds in the androconial extracts was not significantly different between the two groups in any dataset (**Dataset B**: t=1.87, df=5, p=0.12; **Dataset A**: t=1.34, df=4, p=0.26). Therefore, while the number of constituent compounds of the advanced Heliconiini pheromone blend is greater than that of the basal Heliconiini, the distribution of such compounds, as expressed by the evenness,

is similar between the groups, resulting in similar diversity indices. All these results are summed up in Supp. Table 1.4.

An examination of the putative MSPs (present at significantly higher concentrations in the androconia compared to the control tissues) of the Heliconiini reveals a major difference in the pheromonal contents of basal and *Heliconius/Eueides* when it comes to the two broad classes of compounds found in these species: plant and FA derivatives. We refer to all non-FA derived compounds as plant derivatives, as they have all been previously described in plants, though we are not certain whether they are all acquired from plants or other exogenous sources, and lepidopterans are capable of producing plant-like products themselves (Andersson *et al.*, 2003; Estrada *et al.*, 2010; Darragh, Orteu *et al.*, 2019)The compound blends of *Heliconius* and *Eueides* species usually include a mixture of fatty acid and plant derivatives, with the only exception being *H. hierax* (whose only putative MSP is octadecanal, a fatty acid derivative) and most analyzed species in the *sara-sapho* clade of *Heliconius* (whose putative MSPs are all plant derivatives) (Fig. 1.1). By contrast, the putative MSPs of the basal Heliconiini comprise only plant-derived compounds, except for one fatty acid (hexadecanoic acid) found in larger amounts in the androconia of *Agraulis vanillae.* This indicates that while all species are able to produce fatty acid derivatives, the ability to metabolize increased amounts of fatty acid derivatives as part of the androconial blend of compounds is an acquired trait in *Heliconius* and *Eueides*. The basal Heliconiini blend, with its lack of these compounds, appears rather atypical given how widespread fatty acid derivatives are in pheromones blends across lepidopterans (Jurenka and Roelofs, 1993; Tillman et al., 1999; Liénard et al., 2008; Löfstedt et al., 2019). In this sense, FA derivative-inclusive blends can be seen as an ancestral Lepidopteran trait which, in the context of pheromone production, was lost in the basal Heliconiini and subsequently re-gained in *Eueides-Heliconius,*.

Despite the variety of blends observed in the Heliconiini, species remains a very strong predictor of pheromone composition (ADONIS, R^2=0.84, p<0.0001), meaning that androconial contents tend to not vary much within a species. In contrast to species identity, clade identity explains much less variation in pheromone composition (ADONIS, R^2=0.35, p<0.0001). Thus, while there is a tendency of related species to have similar androconial blends, this does not explain much of the diversity seen in the data. The species and clade signatures can be observed in the NMDS plot based on Euclidean distance (that reflects the dissimilarity values utilized by the ADONIS test); conspecifics samples are clustered in this plot, whereas close relatives do not always cluster together, albeit with exceptions such as in the *melpomene-cydno* clade (Supp. Fig. 1.3).

Phylogenetic analysis

Tests on the phylogenetic signal of the full blend on **Dataset B** returned a low value of Kmult (K=0.3).

Similarly low but significant phylogenetic signal was found for all subsets of the data investigated: FA derivatives, plant derivatives and phenolic (lignine) and phytol-derived plant compounds (Table 1.2). This result is discordant with the near absence of phylogenetic signal in *Heliconius* sex pheromones reported previously (Mann et al., 2017). However, the difference is likely due to the more comprehensive sampling of species reported here (33 vs. 11 species analyzed in Mann et al 2017) (Supp. Table 1.1A). Out of all the evolutionary models tested with mvMORPH, the data is best explained by a Brownian motion to early burst (BM-EB) mixed model with the mode shift at the root of *Heliconius* (Table 1.1). Based on the corrected AIC weights, this model is 5.8 times ($AICw_{heliBMEB}/AICw_{advBMEB}=1/0.17$) more likely than the next best model of BM-EB with the shift at the root of (*Eueides* + *Heliconius)* (Wagenmakers and Farrell, 2004) (Table 1.1). This means that a burst of diversification in male sex pheromone composition has accompanied the *Heliconius* radiation. Using the geomorph compare.evol.rates function, we find a significant ~5 fold increase in rate between basal species and *Heliconius* and between *Eueides* and *Heliconius,* but no significant change between basal species and *Eueides*, a result which remains unchanged whether tested on the data itself or on the NMDS axes (on data: net group evolutionary rate $\sigma_B=0.013$, $\sigma_H=0.074$, $\sigma_E=0.222$, $p_{B-H}=0.0001$, $p_{H-E}=0.009$, $p_{B-E}=0.29$; on NMDS axes: $\sigma_B=0.077$, $\sigma_H=0.394$, $\sigma_E=0.142$, $p_{B-H}=0.0001$, $p_{H-E}=0.03$, $p_{B-E}=0.28$).

MOTMOT's unguided tm2 algorithm also finds a shift at the root of *Heliconius* with a ~4 fold increase in evolutionary rate, corroborating what was found with *a priori* information via mvMORPH. However, statistically the rate shift model is not a better fit to the data compared to a simple Brownian model (ΔAICc = 17.36; AICc cutoff = 20.25).

Effect of sympatry on sister species

Based on the assumption that species discriminate conspecifics using differences in pheromone blends, we tested for evidence of character displacement by examining the relationship between the dissimilarity of pheromone blends between 10 Heliconiini sister species pairs and extent of range overlap, while taking account of divergence time. Our linear model, which explains 63% of the variance in pheromone dissimilarity ($R^2=0.63$, $F(7,9)=5.99$, $p=0.03$), shows a significant relationship between sister species pheromone dissimilarity and range overlap (GLM, $\beta=3.59$, $t=2.85$, $p=0.02$), but not with branch length (GLM, $\beta=0.05$, $t=0.48$, $p=0.64$) (Fig. 1.2). Thus, range overlap rather than evolutionary distance more strongly affects differences in pheromones between sister-species pairs.

Correlations among pheromone blend compounds

The correlation matrix (Supp. Fig. 4) shows strong positive associations between several groups of compounds, among which the following are prominent: i) a cluster of phenolic compounds (benzyl salicylate and phenylacetaldehyde) accompanied by their cyanide-containing relatives (benzyl cyanide

and benzyl isocyanide phenylacetaldehyde oxime) and the unrelated (Z)-9-octadecen-1-ol, ii) a subset of potential floral compounds (linalool, cis-linalool oxide in pyranoid form, and methyl salicylate) which may nonetheless be synthesized by the butterflies themselves (Andersson et al., 2003; Darragh, Orteu, et al., 2019) alongside vanillin, iii) several fatty acids, iv) a cluster of straight chain and methylated alkanes, v) several alcohols (including FA-derived alcohols as well as the plant-derived homovanillyl alcohol), vi) phytol-related compounds (including (E)-phytol, (E)-phytal, phytyl acetate and its degradation product hexahydrofarnesylacetone), and vii) phenolic compounds syringic alcohol, acetosyringone and syringic aldehyde. The presence of these clusters of correlated compounds suggests the possibility that compounds within each cluster may share biosynthetic pathways or in the case of chemically unrelated compounds, be required for more effective signalling.

Discussion

We show here that the evolutionarily labile FA metabolic pathway is not deployed in the androconial blends of basal Heliconiini. However, the pathway is reactivated in *Eueides* and *Heliconius* where it is associated with fast diversification of pheromones in the *Heliconius* radiation, in turn making character displacement of chemical signals possible in sympatric sister pairs in relatively short evolutionary times scales. The reactivation of this ubiquitous ancestral pathway seems to have provided the opportunity for niche invasion, with new niches represented by new signal space. Pheromone dissimilarity between sympatric pairs is dramatic, with the two species often having no major compounds in common, a testament to the rapidity with which these signals can evolve. These chemical blends shown a high degree of species-specificity and usually very low within-species variation, both important properties of traits involved in assortative mating (Cardé et al., 1977; Cardé, 1986; Smadja and Butlin, 2009). These results provide the strongest evidence to date of the role of male sex pheromones as reproductive barriers across the *Heliconius* radiation.

FA derivatives-driven burst in the diversification of *Heliconius* pheromones

Phylogenetic analysis of our data suggest that factors other than neutral change (Losos, 2008; Harmon, 2019) are responsible for the evolution of the wide variety of pheromone blends among the Heliconiini. A value of $K<1$ indicates that the traits are more diversified among close relatives than they would be under random expectations (Blomberg et al., 2003; Revell et al., 2008). With little dependence on their phylogenetic history, these traits could be considered evolutionarily labile and thus able to diversify rapidly, a pattern often seen in behavioural traits involved in intraspecific signalling such as bird songs (Blomberg et al., 2003; Rheindt et al., 2004). In all cases, FA derivatives showed weaker phylogenetic signal than plant derivatives (Table 1.2), indicating that while both compound classes are evolutionarily labile, the former is relatively more flexible. The higher K value

obtained when including only advanced species is more likely due to reduced resolution of the test, than to high diversity of the basal species.

We observe a burst in pheromone diversification rate at the root of *Heliconius*. Given the striking difference in androconial composition between the basal (predominantly plant derivatives) and *Eueides/Heliconius* (mixtures of fatty acid and plant derivatives) species, it appears that re-activation of the FA metabolic pathway for pheromone production triggered this diversification. Obtaining the metabolic or physiological tools to invade new niche spaces is often accompanied by rapid diversification as seen in such examples as corallivory in *Chaetodon* butterflyfishes (Bellwood et al., 2010; Konow et al., 2017) or the shift to a nectarivorous diet in Hawaiian honeycreepers (Lovette et al., 2002). The *Heliconius* radiation itself is known to have probably been facilitated by a key innovation not observed in any other Lepidopteran: pollen feeding, which brought about an array of new adaptations (Gilbert, 1972, 1975; Young and Montgomery, 2020). While the FA metabolic pathway re-activation constitutes a less dramatic physiological change, it still had an analogous effect in making new chemical sexual signalling tools available to *Heliconius* and *Eueides.*

Somewhat surprisingly, *Eueides* species use FA products in their pheromone blends yet do not show the same blend diversity as *Heliconius*. This may be an artefact of our dataset containing only 3 of the 9 described species of *Eueides* (Kozak *et al.*, 2015). It is plausible that with the inclusion of more *Eueides* species, the evolutionary shift may indeed be detected at the root of *Heliconius/Eueides*, rather than the root of just *Heliconius.*

In spite of their absence in the pheromone blends of the basal Heliconiini, FA-derived compounds are nearly ubiquitous in Lepidopteran pheromones (Jurenka and Roelofs, 1993; Tillman et al., 1999; Liénard et al., 2008), and there is some evidence that the biosynthetic pathways are conserved between moths and butterflies (Liénard et al., 2014). *Cethosia*, often considered part of a sister group to the Heliconiini tribe (Freitas and Brown, 2004; Silva-Brandão et al., 2008), produces chemical signals that include both plant derivatives and FA derivatives (Li et al., 2017). This indicates that the FA pathway still retained its role in pheromone production at the time of the *Cethosia*-Heliconiini split. Therefore, FA-based pheromones appear to be an ancestral Lepidopteran trait, which was subsequently lost at the root of the Heliconiini and regained with the split of *Eueides/Heliconius* from the basal genera.

In addition to the clear differences in pheromone composition between basal and advanced species, with our comprehensive sampling of species we find that while no compound is exclusive to a single species, each species has its own unique blend of compounds, confirming previous work carried out on fewer species (Mann et al., 2017; Darragh et al., 2020). The high observed species specificity of pheromone composition is consistent with the role of pheromones as potential reproductive barriers

in these species, and it is partially achieved through the implementation of the FA pathway, making it one more consequence of its reactivation in *Eueides*/*Heliconius*. Traits involved in assortative mating are known to evolve under divergent selection (Bolnick and Fitzpatrick, 2007; Merrill et al., 2012) yet strong species-specificity of the mating cue is expected, especially in species with overlapping distributions (Andersson et al., 2007; Smadja and Butlin, 2009), in order to obtain a response and provide reliable information to the receiving party (Andersson et al., 2007).

Sympatry increases pheromone differentiation in sister species

Our findings that pheromone differentiation increases with the extent of range overlap between sister species provides by far the most compelling evidence that pheromone composition is important for reproductive isolation among heliconiine species. This pattern of increased divergence between sympatric pairs is consistent with a scenario of reinforcement. Reinforcement is a phenomenon that may be observed when speciation occurs in complete or partial sympatry, where gene flow between the incompletely reproductively isolated taxa may promote the evolution of reproductive isolation because of selection for increased mate discrimination (Dieckmann and Doebeli, 1999). An important test of reinforcement is the relative strength of pre-zygotic reproductive barriers between sister taxa in different geographic contexts, with stronger barriers expected between sympatric taxa, with the potential for gene flow, than allopatric taxa (Noor, 1999; Servedio and Noor, 2003; Ortiz-Barrientos et al., 2009). This condition is unquestionably fulfilled in Heliconiini pheromones, where the strikingly divergent compositions seen in sympatric pairs would greatly improve discrimination in favour of conspecifics.

Since scent is a trait whose divergence reduces the likelihood of hybridization, pheromone divergence would be considered a case of reproductive character displacement (Noor, 1999; Servedio and Noor, 2003; Pfennig and Pfennig, 2009). Similar pheromone patterns are seen in *Bicyclus* (Bacquet et al., 2015), orchid bee pheromones (Weber et al., 2016), *Hemileuca* moths (McElfresh and Millar, 2001), and in the Tortricidae (Cardé et al., 1977; Hill et al., 1977). It is however far from universally observed in sympatric species of insects (Symonds and Elgar, 2004), perhaps due to the varying importance of sex and aggregation pheromones in reproduction compared to other factors such as host plant or habitat preference.

Previous studies have highlighted that mate choice in many species of *Heliconius* butterflies is based on colour patterns, with males showing assortative mating based on visual cues (Jiggins *et al.*, 2001; Naisbit, Jiggins and Mallet, 2001; Melo *et al.*, 2009; Merrill, Chia and Nadeau, 2014; Merrill *et al.*, 2015; Rosser *et al.*, 2019). However colour alone is not sufficient to explain the extent to which males discriminate towards heterospecific live females (Jiggins *et al.*, 2001), or between closely-related co-mimetic species (Mérot et al., 2015), implying that more signals may be at play. Here we provide

evidence that male sex pheromones are important for initiating and/or maintaining species reproductive barriers across this group of butterflies.

There are two main implications to our findings: 1) that the re-activation of the FA metabolic pathway, originally lost in the pheromone-producing tissues of basal Heliconiini, accompanied the increased rate of speciation seen in this genus compared to other Heliconiini, meaning that it unlocked ecological opportunity. 2) that male sex pheromones show macroevolutionary patterns expected from pre-zygotic reproductive barriers, with strong species-specificity, character displacement and the ability to diverge quickly in closely related species when needed, all of which minimize the chance of hybridization. The recruitment of a single family of enzymes, the Δ-11-desaturases, while themselves seemingly missing in *Heliconius* in favour of the widespread Δ-9-desaturases (Mann *et al.*, 2017) preceded the Lepidopteran radiation, and is known to be key for pheromone diversification in the entire order, making numerous compound families available as pheromone components (Liénard *et al.*, 2008). What happened with the FA pathway at the root of *Eueides*/*Heliconius* may be analogous to this, though at a smaller scale. It is not known whether pheromone diversification has been a primary driver for the iconic *Heliconius* radiation, perhaps concurrently with ecological differences such as habitat or host plant preferences, or whether it reinforced mating isolation following the diversification of other courtship cues such as the colour pattern. Nonetheless, we have detected, for the first time, strong evidence that chemical signalling cues have had an important part in the evolution and maintenance of the many *Heliconius* species seen today.

Figures and Tables

Figure 1.1. *(below)*: Heliconiini tree with branches coloured according to disparity in pheromone profiles. Values for disparity (expressed by Euclidean distance) are shown at each node. The table shows putative male sex pheromones composition (**dataset C**) of each species, where columns represent androconia-exclusive compounds and amounts are expressed as the proportion of the total androconial contents. Note the near absence fatty acid derivatives among the basal genera, while in general *Eueides* and *Heliconius* species produce a varied cocktail comprising both plant and fatty acid derivatives. This figure summarises the most abundant compounds (at least 12% of the total androconial contents). For a complete dataset with all compounds, see (Supp. Fig. 1.5)

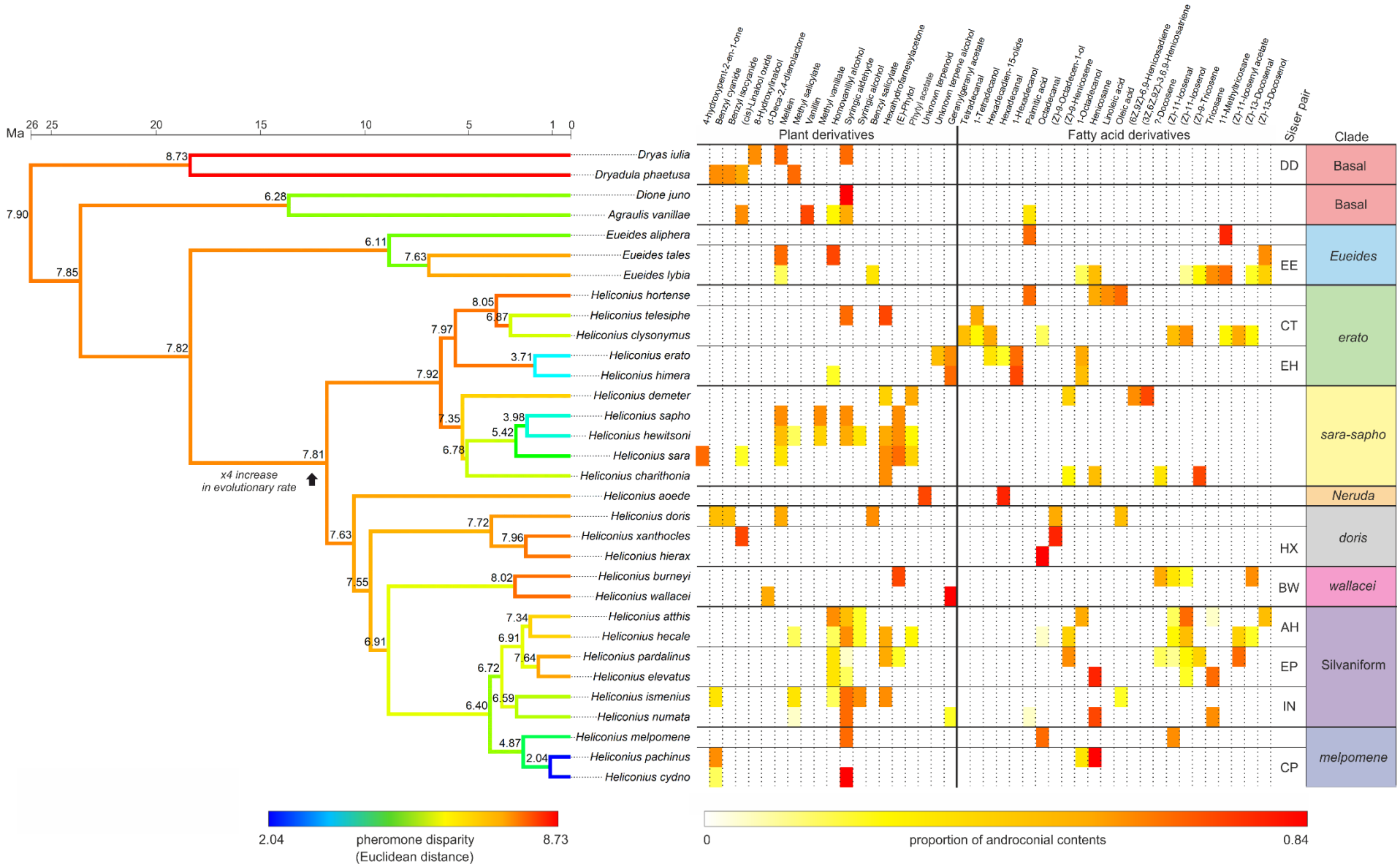

Ma
26
25
20
15
10
5
1
0
8.73
6.28
7.90
6.11
7.63
7.85
8.05
6.87
7.97
7.82
3.71
7.92
7.35
3.98
5.42
6.78
7.81
x4 increase
in evolutionary rate
7.72
7.96
7.63
7.55
8.02
7.34
6.91
6.91
7.64
6.72
6.59
6.40
4.87
2.04
Dryas iulia
Dryadula phaetusa
Dione juno
Agraulis vanillae
Eueides aliphera
Eueides tales
Eueides lybia
Heliconius hortense
Heliconius telesiphe
Heliconius clysonymus
Heliconius erato
Heliconius himera
Heliconius demeter
Heliconius sapho
Heliconius hewitsoni
Heliconius sara
Heliconius charithonia
Heliconius aoede
Heliconius doris
Heliconius xanthocles
Heliconius hierax
Heliconius burneyi
Heliconius wallacei
Heliconius atthis
Heliconius hecale
Heliconius pardalinus
Heliconius elevatus
Heliconius ismenius
Heliconius numata
Heliconius melpomene
Heliconius pachinus
Heliconius cydno
4-hydroxypent-2-en-1-one
Benzyl cyanide
Benzyl isocyanide
(cis)-Linalool oxide
8-Hydroxylinalool
α-Deca-2,4-dienolactone
Mellein
Methyl salicylate
Vanillin
Methyl vanillate
Homovanillyl alcohol
Syringic aldehyde
Benzyl alcohol
Benzyl salicylate
Hexahydrofarnesylacetone
(E)-Phytol
Phytyl acetate
Unknown terpenoid
Unknown terpene alcohol
Geranylgeranyl acetate
Plant derivatives
1-Tetradecanol
Tetradecanal
Hexadecadien-15-olide
1-Hexadecanol
Hexadecanal
Palmitic acid
Octadecanal
(Z)-9-Octadecen-1-ol
(Z)-9-Henicosene
1-Octadecanol
Henicosane
Linoleic acid
Oleic acid
(6Z,9Z)-6,9-Henicosadiene
(3Z,6Z,9Z)-3,6,9-Henicosatriene
?-Docosene
(Z)-11-Icosenal
(Z)-11-Icosenol
(Z)-9-Tricosene
Tricosane
11-Methyltricosane
(Z)-11-Icosenyl acetate
(Z)-13-Docosenal
(Z)-13-Docosenol
Fatty acid derivatives
Sister pair
Clade
DD
Basal
Basal
EE
Eueides
CT
EH
erato
sara-sapho
Neruda
HX
doris
BW
wallacei
AH
EP
IN
Silvaniform
CP
melpomene
2.04
pheromone disparity
(Euclidean distance)
8.73
0
proportion of androconial contents
0.84

Table 1.1. Evolutionary models of Heliconiini male sex pheromone diversification ordered from best (top row) to worst fit (bottom row). The best fitting model starts with Brownian motion followed by a burst of diversification at the root of *Heliconius.* The corrected Akaike Information Criterion (AICc) was used to assess model fit, and AICw (AIC weight) represents the likelihood of a particular model being the best fitting one among the 6 tested models.

Model description	AICc	AIC diff	AICw	mode	N shifts	shift location
Brownian to early burst	603	0.	1.00	BM-EB	1	*Heliconius*
Brownian to early burst	607	3.56	0.17	BM-EB	1	*Eueides-Heliconius*
Brownian with a rate shift	611	7.42	0.02	BMM	1	*Heliconius*
Brownian	622	18.57	0.00	BM	0	-
Early burst	624	21.23	0.00	EB	0	-
Evolutionary constraint	629	25.53	0.00	OU	0	-

Figure 1.2. Relationship between dissimilarity in male sex pheromones (Euclidean distance) and range overlap between heliconiine sister species pairs. Divergence time between each sister species pair is represented as a colour scale. Each point is a pair of sister species: AH = *H. atthis-H. hecale*, BW = *H. burneyi-H. wallacei*, CP = *H. cydno-H. pachinus*, CT = *H. clysonimus-H. telesiphe*, DD = *Dryas iulia-Dryadula phaetusa*, EE = *Eueides tales-Eueides lybia*, EH = *H. erato-H. himera*, EP = *H. elevatus- H. pardalinus*, HX = *H. hierax-H. xanthocles*, IN = *H. ismenius-H. numata*.

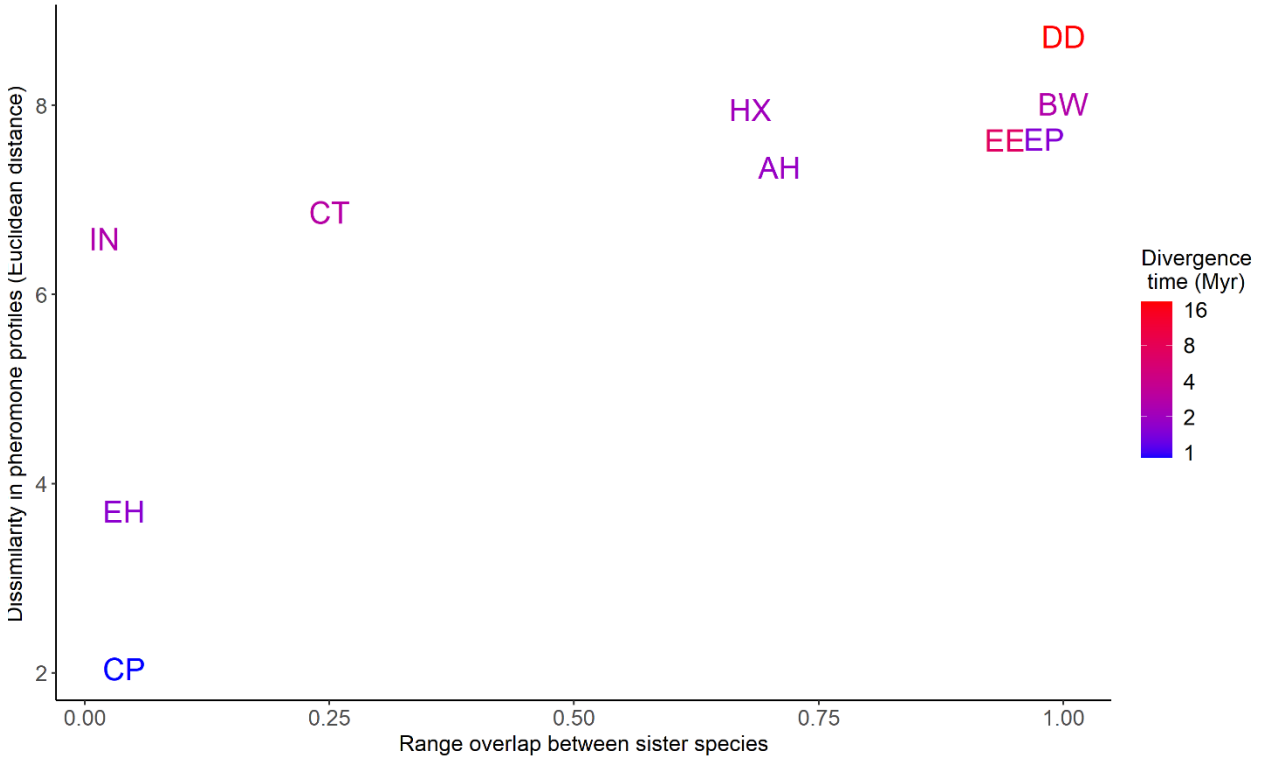
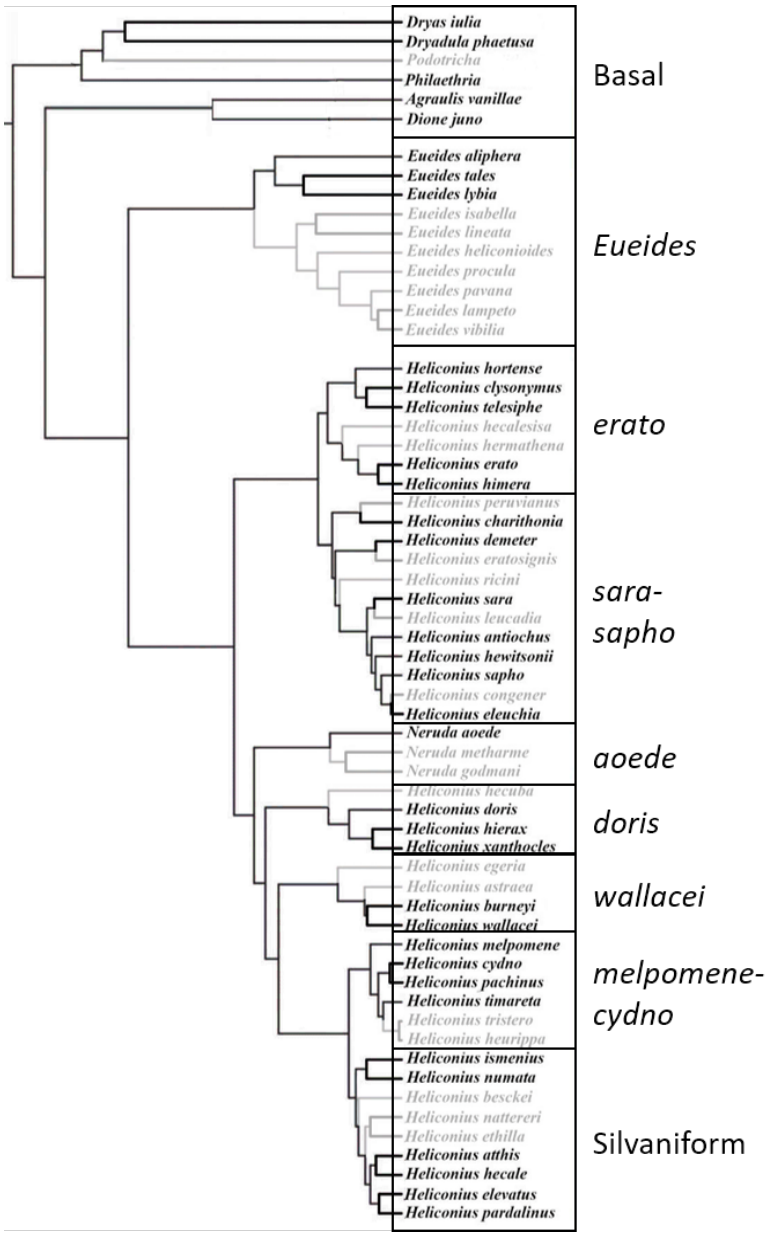


Table 1.2. Phylogenetic signal (Kmult) calculated for different subsets of androconia compounds. Kmult measures phylogenetic signal of a trait in relation to a Brownian motion (random walk) evolutionary model. K=1 or close to that value describes a Brownian signal, with closely related species being similar; K>1, never observed in this data, denotes greater phylogenetic signal and thus greater evolutionary constraint than expected under Brownian assumptions; K<1 denotes lower signal than expected under Brownian assumptions, with K=0 representing an absence of signal. Note that due to the small number of basal species in the study, Kmult values for this group are never significant.
**These compounds are not found in all species. **These compounds are not found in any species of this group.*

	Dataset B			Confidence interval	
Compounds	Subset	Kmult	p-value	5%	95%
All compounds	All species	0.30	**0.0001**	0.18	0.22
	Eueides-Heliconius	0.40	**0.0001**	0.25	0.29
	Basal	0.91	0.09	0.78	0.91
Fatty acid derivatives	All species	0.27	**0.0001**	0.18	0.22
	Eueides-Heliconius	0.36	**0.0001**	0.25	0.29
	Basal	1.04	0.23	0.74	1.07
Plant derivatives	All species	0.32	**0.0001**	0.18	0.22
	Eueides-Heliconius	0.44	**0.0001**	0.24	0.29
	Basal	0.82	0.42	0.76	1.
Phytol	All species*	0.29	0.1127	0.2	0.31
	Eueides-Heliconius *	0.39	0.0517	0.26	0.39
	Basal**	-	-	-	-
Lignin phenols	All species	0.32	**0.0001**	0.17	0.23
	Eueides-Heliconius	0.47	**0.0001**	0.23	0.3
	Basal	0.78	0.77	0.7	1.1
Phytol + phenols	All species	0.33	**0.0001**	0.18	0.22
	Eueides-Heliconius	0.46	**0.0001**	0.24	0.3
	Basal	0.78	0.75	0.71	1.06

Supplementary Material

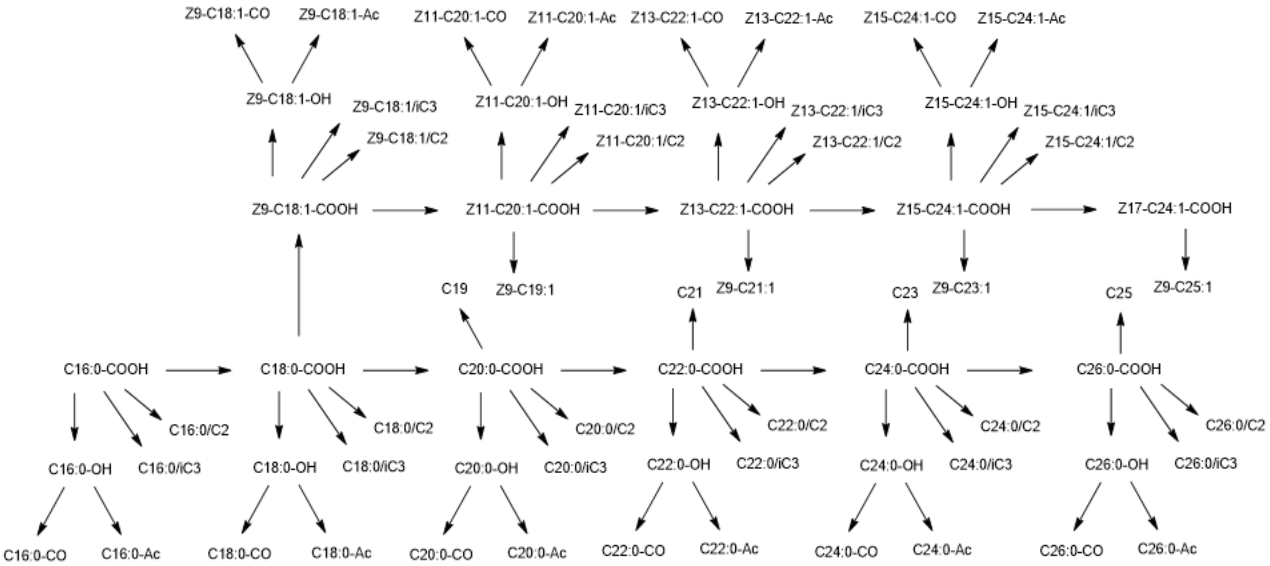
A.



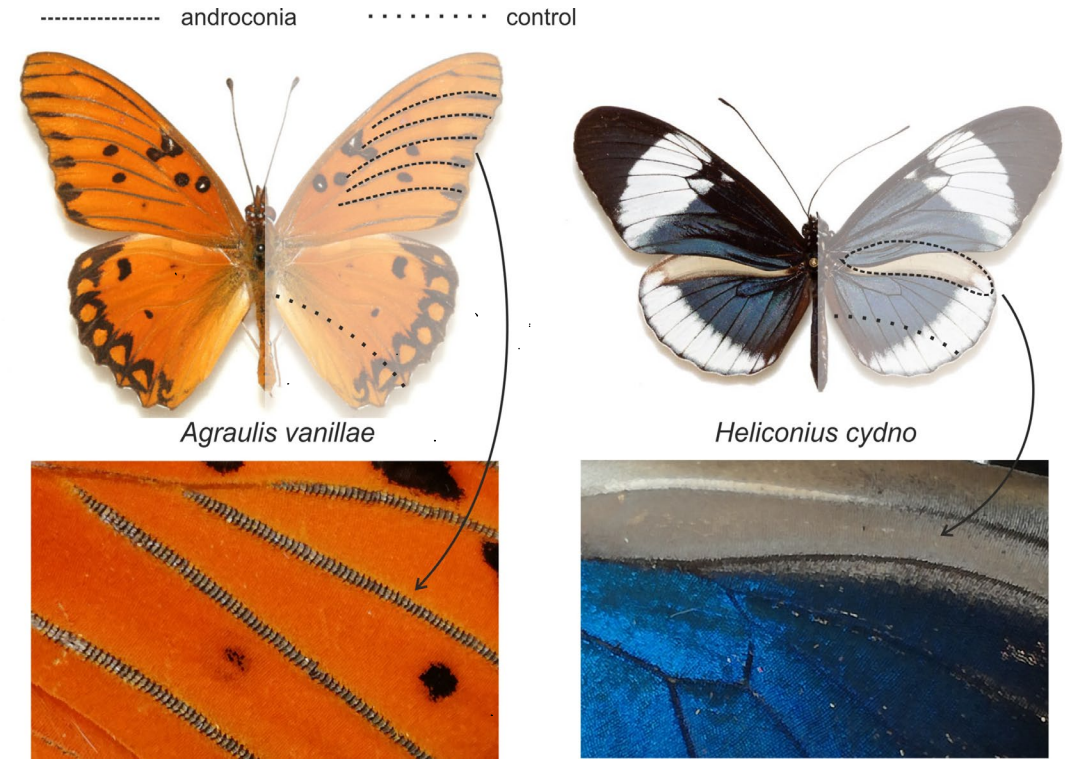
B.

Species	N. samples		
	male	female	control
Agraulis vanillae	5	1	5
Dione juno	4	2	4
Dryadula phaetusa	4	0	4
Dryas iulia	8	2	8
Eueides aliphera	6	2	6
Eueides lybia	8	1	8
Eueides tales	2	0	2
H. antiochus	1	0	1
H. aoede	3	3	3
H. atthis	5	6	5
H. burneyi	1	0	1
H. charithonia	5	6	5
H. clysonimus	4	0	4
H. cydno	4	1	4
H. demeter	3	1	3
H. doris	4	3	4
H. eleuchia	9	0	0
H. elevatus	9	7	9
H. erato	1	0	1
H. erato etylus	3	0	3
H. erato notabilis	3	0	3
H. hecale	5	2	5
H. hewitsonii	7	2	7
H. hierax	2	0	2
H. himera	2	0	2
H. hortense	4	3	4
H. ismenius	7	5	7
H. melpomene	6	0	0
H. melpomene plesseni	4	0	4
H. numata	4	2	4
H. pachinus	2	1	2
H. pardalinus butleri	10	7	10
H. pardalinus sergestus	10	5	9
H. sapho	2	0	2
H. sara	3	3	3
H. telesiphe	3	0	3
H. timareta	8	0	0
H. wallacei	4	1	4
H. xanthocles	3	0	3
Philaethria dido	1	0	1

Supp. Table 1.1. A:Heliconiini phylogeny. Species featured in the study are shown in black, and the rectangles highlight the different subclades as well as the basal genera (truly a polyphyletic clade). B: Number of samples collected for each species.



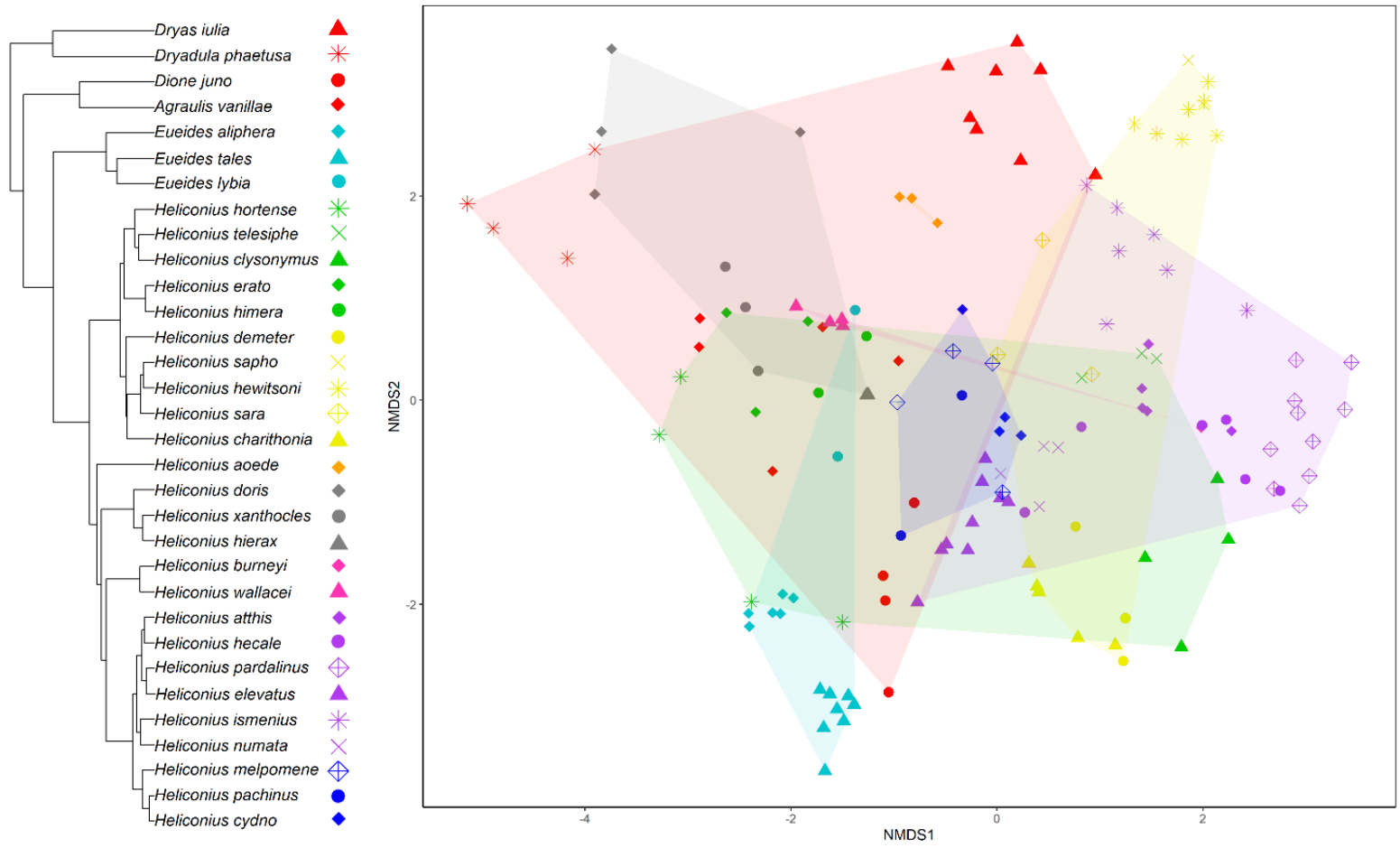
Supp. Figure 1.1. The FA metabolic pathway described by Mann et al. (2017). -COOH=fatty acids, -OH= alcohols, -CO=aldehydes, -Ac=esters. Cn:m determines the position and number of double bonds. Cn:0=saturated (no double bond), Cn:1=unsaturated (single double bond).



Supp. Figure 1.2. Androconia and control regions extracted from basal and advanced genera, represented here by *A. vanillae* and *H. cydno* respectively, alongside closeup pictures of the two different androconial configurations. Pinned specimen pictures courtesy of https://butterfliesofamerica.com/, credited to A. D. Warren (2010) and K. Davis, M. Stangeland & A. Warren (2009) respectively.

Species	race	Clade	Region	Collection site	Range size (km²)	Habitat	Openness	Mimicry	N males	N females
Dryas iulia		Basal	Panama	STRI Gamboa, Panama	15996690.99	Field	4	Orange	9	2
Dryadula phaetusa		Basal	Panama	STRI Gamboa, Panama	13245381.63	Field	4	Orange	4	
Philaethria dido		Basal	Panama	Pipeline Road, Gamboa, Panama		Closed forest	1	Green	1	1
Dione juno		Basal	Panama	STRI Gamboa, Panama	14769311.44	Forest	2	Orange	4	2
Agraulis vanillae		Basal	Panama	STRI Gamboa, Panama	22164407.64	Field	4	Orange	5	1
Eueides aliphera	*gracilis*	Eueides	Panama	Pipeline Road, Gamboa, Panama	14118176.99	Edge	3	Orange	6	2
Eueides tales	*calathus*	Eueides	Ecuador	Campococha, Ecuador	5979225.28	Tall forest	2	Rays	3	
Eueides lybia	*olympia*	Eueides	Panama	Pipeline Road, Gamboa, Panama	6580704.98	Gaps	3	Orange	8	1
Heliconius clysonymus	*montanus*	Erato	Panama/Ecuador	Mount Totumas, Panama/Mina Negra, Ecuador	304800.97	Cloud forest	1	RW	1	1
Heliconius telesiphe	*sotericus*	Erato	Ecuador	Various	247235.21	Cloud forest	1	Postman*	3	
Heliconius hortense		Erato	El Salvador	Captive	579540.43	Cloud forest	1	RW	4	3
Heliconius erato	*demophoon*	Erato	Panama	STRI Gamboa, Panama		Edge	3	Postman	1	
Heliconius erato	*notabilis*	Erato	Ecuador	Captive	11728.51	Edge	3	Dots	3	
Heliconius erato	*etylus*	Erato	Ecuador	Captive	36273.68	Edge	3	Rays	3	
Heliconius himera		Erato	Ecuador	Captive	38806.31	Scrub	4	RW	2	
Heliconius sapho	*sapho*	Sara-sapho	Panama	Pipeline Road, Gamboa, Panama	787074.18	Closed forest	1	Blue	2	
Heliconius eleuchia	*primularis*	Sara-sapho	Ecuador	West Ecuador	548590.29	Riparian forest	1	Rays	9	
Heliconius hewitsoni		Sara-sapho	Panama	El Montuoso, Panama	21219.00	Forest	2	Blue	7	2
Heliconius antiochus	*araneides*	Sara-sapho	Suriname	Captive		Riparian forest	1	Blue	1	
Heliconius sara	*magdalena*	Sara-sapho	Panama	STRI Gamboa, Panama	13076244.86	Edge/Gaps	3	Blue	3	3
Heliconius demeter	*neildi*	Sara-sapho	Ecuador	Campococha, Ecuador		Sand forest	3			
Heliconius charithonia		Sara-sapho	Various	Captive	4985381.27	Edge/Gaps	3	BW	5	6
Heliconius aoede	*aliciae*	Neruda	Ecuador	Campococha, Ecuador	6772.99	Closed forest	1	Rays	10	
Heliconius hierax		Doris	Ecuador	Various	35462.41	Cloud forest	1	RW	3	
Heliconius xanthocles	*napoensis*	Doris	Ecuador	Various	7278498.78	Marginal forest	2	Rays	3	
Heliconius doris		Doris	Panama	STRI Gamboa, Panama	7235981.23	Gaps	3	Mixed	4	2
Heliconius burneyi	*huebneri*	Wallacei	Ecuador	Campococha, Ecuador		Tall forest	2	Rays	1	
Heliconius wallacei	*flavescens*	Wallacei	Ecuador	IKIAM campus, Ecuador	7548196.03	Tall forest	2	Blue	5	1
Heliconius atthis		Silvaniform	Ecuador	Captive	84725.44	Marginal forest	2	BW	5	6
Heliconius hecale	*melicerta*	Silvaniform	Panama	Pipeline Road, Gamboa, Panama	7966004.74	Forest	2	Silvaniform	5	2
Heliconius pardalinus	*butleri*	Silvaniform	Peru	Tarapoto, Peru	559092.32	Riparian forest	1	Silvaniform	13	2
Heliconius pardalinus	*sergestus*	Silvaniform	Peru	Tarapoto, Peru	5712.22	Riparian forest	1	Silvaniform	5	
Heliconius elevatus		Silvaniform	Peru	Tarapoto, Peru	5169877.39	Riparian forest	1	Rays	10	2
Heliconius ismenius		Silvaniform	Ecuador/Mixed	Captive	1187146.85	Forest	2	Silvaniform	8	5
Heliconius numata	*tarapotensis*	Silvaniform	Peru	Tarapoto, Peru	11491367.93	Forest	2	Silvaniform	4	3
Heliconius melpomene	*rosina*	Cydno	Panama	STRI Gamboa, Panama	171722.08	Edge	3	Postman	6	
Heliconius melpomene	*plesseni*	Cydno	Ecuador	Captive	10758.30	Edge	3	Dots	4	
Heliconius timareta		Cydno	Ecuador	Colombia, West Ecuador	34394.88	Montane forest	1	Rays	8	
Heliconius pachinus		Cydno	Panama	El Montuoso, Panama	26793.15	Forest	2	Blue	2	1
Heliconius cydno	*chioneus*	Cydno	Panama	Pipeline Road, Gamboa, Panama	601073.76	Edge/Gaps	3	Blue	4	1

Supp. Table 1.2. General information on all Heliconiini species featured in the study. Range size was calculated by Neil Rosser (Harvard University) based on data published in Rosser *et al.*, 2012. Habitat data from (Brown, 1981) with openness and mimicry ring from Queste (2020).



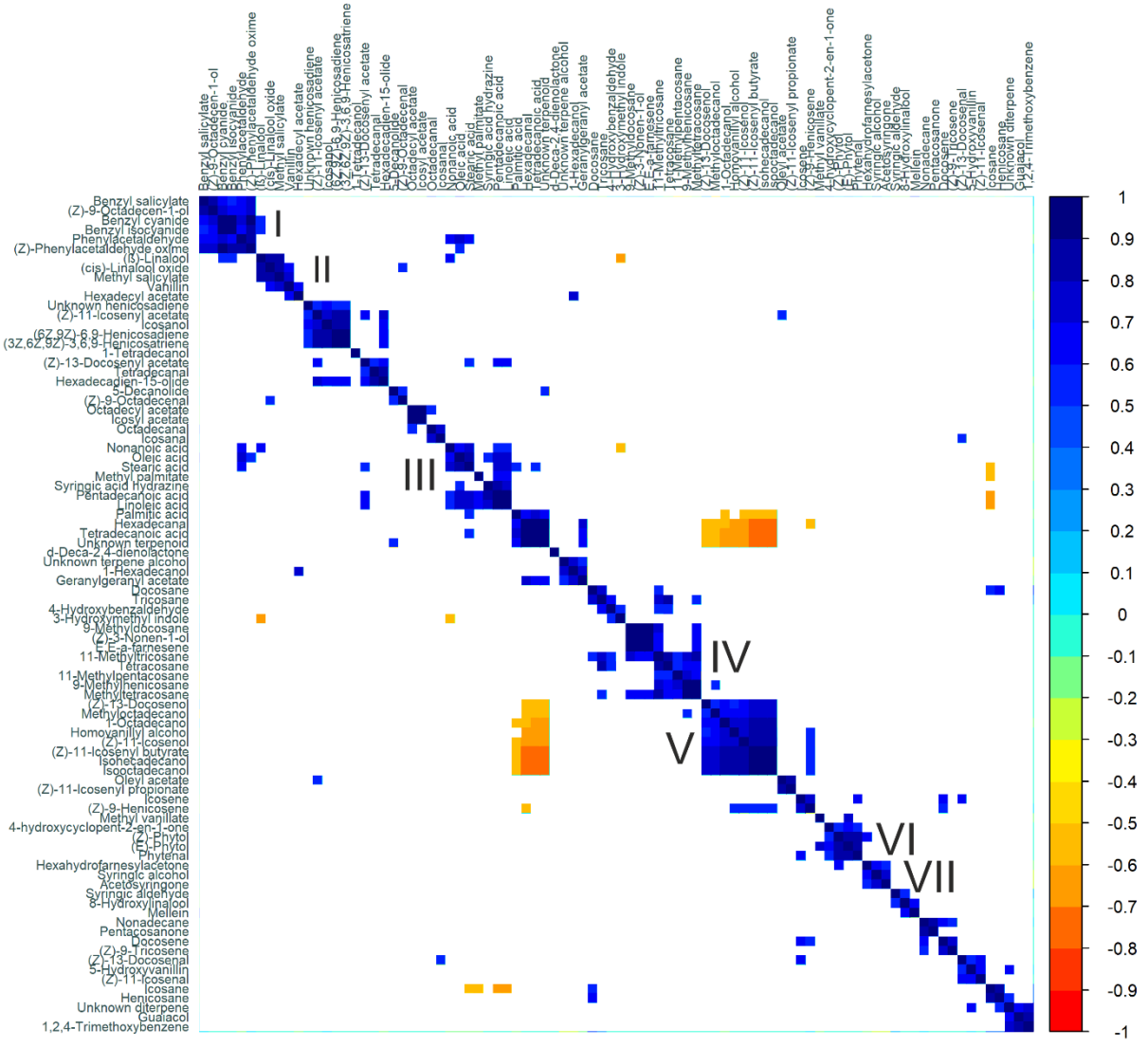
Supp. Figure 1.3. NMDS based on Euclidean distances within the filtered dataset. Hulls encompass the clades. This plot excludes *H. timareta* and *H. eleuchia* as no controls were available to apply any filtering to those species. Extra species are also not included.

	Total amount in 200µl (nmol)			Avg. number of compounds			Diversity (Shannon index)			Evenness			N. t-test sig. compounds
Species	M	F	C	M	F	C	M	F	C	M	F	C	
Agraulis vanillae	**5.05**	1.15	2.94	**15.2**	7.0	10.0	**1.64**	0.53	0.94	**0.55**	0.21	0.34	7
Dione juno	**10.39**	4.37	10.11	13.75	**17.5**	15.0	**1.33**	1.02	1.31	**0.45**	0.33	0.44	1
Dryadula phaetusa	**4.39**	*NA*	4.27	**12.75**	*NA*	12.25	**0.44**	*NA*	0.33	**0.15**	*NA*	0.11	4
Dryas iulia	**5.57**	3.24	2.48	8.5	**17.5**	6.5	1.24	**1.35**	0.44	**0.48**	0.44	0.18	4
Eueides aliphera	**29.5**	1.21	14.77	**17.67**	8.5	14.0	1.15	**0.61**	1.06	**0.37**	0.23	0.36	5
Eueides lybia	**28.69**	2.11	4.0	**26.5**	7.0	15.63	**1.75**	0.27	1.24	**0.51**	0.11	0.41	20
Eueides tales	9.75	*NA*	**45.36**	**9.0**	*NA*	7.0	**0.72**	*NA*	0.46	**0.27**	*NA*	0.18	3
H. aoede	**53.14**	3.43	17.55	19.33	19.0	**24.67**	**1.78**	1.72	1.22	**0.56**	0.54	0.36	4
H. atthis	**18.59**	1.89	1.22	**31.2**	13.0	11.6	**1.83**	0.98	0.43	**0.51**	0.34	0.15	15
H. burneyi	**75.5**	*NA*	3.23	**24.0**	*NA*	0.0	1.65	*NA*	**1.97**	0.49	*NA*	**0.58**	-
H. charithonia	**12.45**	2.16	2.48	**23.4**	17.0	12.2	1.39	**0.42**	0.58	**0.42**	0.13	0.19	7
H. clysonimus	**77.88**	*NA*	11.65	**32.75**	*NA*	18.0	**2.19**	*NA*	1.23	**0.6**	*NA*	0.39	13
H. cydno	**16.59**	1.28	2.02	**22.5**	21.0	10.0	1.3	**0.97**	0.94	**0.39**	0.3	0.35	5
H. demeter	**45.7**	4.82	5.25	**28.33**	15.0	16.67	**1.41**	1.23	1.35	0.4	0.4	**0.43**	6
H. doris	**9.09**	1.31	1.12	16.5	**22.0**	8.5	**1.92**	0.99	0.47	**0.63**	0.29	0.18	8
H. eleuchia	**9.86**	*NA*	*NA*	**10.22**	*NA*	*NA*	**1.15**	*NA*	*NA*	**0.42**	*NA*	*NA*	-
H. elevatus	**30.81**	2.41	4.05	**20.22**	8.43	11.67	**1.21**	0.59	1.21	0.38	0.22	**0.43**	4
H. erato etylus	**2.91**	*NA*	1.03	**11.67**	*NA*	6.67	**1.52**	*NA*	0.14	**0.54**	*NA*	0.06	5
H. erato notabilis	**6.13**	*NA*	1.04	**17.67**	*NA*	7.67	**1.69**	*NA*	0.21	**0.54**	*NA*	0.08	6
H. hecale	**11.63**	3.21	3.38	**26.4**	16.5	6.4	**1.8**	0.38	0.39	**0.52**	0.12	0.16	19
H. hewitsonii	**12.85**	5.37	5.56	**20.29**	10.5	11.57	**1.78**	0.75	0.71	**0.55**	0.27	0.25	10
H. hierax	**102.08**	*NA*	4.2	16.5	*NA*	**20.5**	0.78	*NA*	**1.68**	0.26	*NA*	**0.52**	2
H. himera	**4.46**	*NA*	3.02	**16.5**	*NA*	7.5	**0.92**	*NA*	0.05	0.3	*NA*	0.02	5
H. hortense	**10.27**	4.13	4.21	**28.5**	14.67	13.0	**1.55**	0.21	0.26	**0.44**	0.07	0.09	8
H. ismenius	**5.57**	4.34	4.17	**19.43**	16.2	8.86	**0.93**	0.39	0.2	0.29	0.13	0.08	8
H. melpomene plesseni	**14.57**	*NA*	6.27	**20.75**	*NA*	10.75	**1.38**	*NA*	0.21	**0.42**	*NA*	0.07	5
H. numata	**37.64**	5.66	10.52	**26.5**	12.0	16.5	**1.39**	0.49	1.1	0.4	0.17	0.35	15
H. pachinus	**13.13**	3.42	8.72	**21.5**	11.0	16.0	**1.35**	0.55	0.9	**0.41**	0.2	0.28	3
H. pardalinus butleri	**44.56**	6.09	6.24	**25.5**	6.71	7.8	**1.72**	0.1	0.2	**0.51**	0.04	0.08	16
H. sapho	**17.3**	*NA*	5.02	**15.0**	*NA*	6.5	**1.72**	*NA*	0.03	**0.58**	*NA*	0.01	4
H. sara	**15.91**	5.19	5.06	**21.0**	10.33	8.33	**1.54**	0.19	0.08	**0.47**	0.06	0.03	7
H. telesiphe	**11.65**	*NA*	8.79	**14.33**	*NA*	12.33	**1.25**	*NA*	**0.85**	**0.42**	*NA*	0.29	3
H. timareta	**10.13**	*NA*	*NA*	**12.5**	*NA*	*NA*	**0.97**	*NA*	*NA*	**0.34**	*NA*	*NA*	-
H. wallacei	9.38	2.38	**17.51**	8.25	7.0	**11.5**	**0.75**	**0.6**	0.68	**0.29**	0.24	0.24	2
H. xanthocles	**4.43**	*NA*	3.32	11.0	*NA*	**12.0**	**0.86**	*NA*	**0.47**	**0.31**	*NA*	0.17	2
Philaethria dido	**19.36**	*NA*	8.49	**14.0**	*NA*	0.0	**0.98**	*NA*	0.94	0.33	NA	**0.33**	-

Supp. Table 1.3. Comparison between the contents of male androconia, female "androconia" and control regions of the wing, including the total concentration of androconial contents, average number of compounds (richness) diversity and evenness of compounds for each species, and number of compounds that showed significant differences in amount between androconia and control (N. t-test sig. compounds).

		Welch's two-sample t-test			*Eueides-Heliconius*		Basal		
Parameter	Dataset	t	df	*p*- value	mean	SD	mean	SD	mean diff.
Diversity of compounds	A	1.34	3.82	*0.26*	2.28	0.38	2.01	0.38	0.27
	B	2.25	4.41	*0.08*	2.11	0.38	1.74	0.30	0.37
	C	1.87	4.89	*0.12*	1.41	0.54	1.01	0.38	0.40
Richness (N compounds)	A	4.36	8.86	***0.002***	20.25	7.57	11.26	3.10	8.99
	B	4.13	7.73	***0.004***	17.16	6.30	9.55	2.88	7.61
	C	3.09	8.12	***0.01***	6.89	4.23	3.10	1.90	3.80
Evenness of compounds	A	-2.36	3.65	*0.08*	0.78	0.05	0.86	0.06	-0.08
	B	-1.30	3.70	*0.27*	0.77	0.06	0.81	0.06	-0.04
	C	0.80	3.04	*0.48*	0.79	0.17	0.47	0.79	0.32

Supp. Table 1.4. Results of Welch's two-sample t-test between basal and advanced species for diversity, richness and evenness of the pheromone blend, and for all levels of filtering of the data: unfiltered, filtered by including only compounds that were significantly more abundant in the androconia of any species, and lastly by removing from each species all non-significant compounds, leaving only putative MSPs.



Supp. Figure 1.4. Correlation matrix between compounds in the filtered dataset, excluding compounds that appeared a single time. Each tile represents a correlation test between two compounds. Colour indicates the strength of the correlation (Pearson's correlation coefficient). Blank tiles indicate non-significant results.

Sister species pair	Dissimilarity (Euclidean distance)	Range overlap	Branch length (Myr)
H. cydno-H. pachinus	2.04	0.04	0.9
H. erato-H. himera	3.71	0.04	1.65
H. ismenius-H. numata	6.59	0.02	2.54
H. clysonimus-H. telesiphe	6.87	0.25	2.84
H. atthis-H. hecale	7.34	0.71	1.87
E. lybia-E. aliphera	7.63	0.94	6.81
H. elevatus-H. pardalinus	7.64	0.98	1.51
H. hierax-H. xanthocles	7.96	0.68	2.09
H. burneyi-H. wallacei	8.02	1	2.62
Dryas iulia-Dryadula phaetusa	8.73	1	18.44

Supp. Table 1.5. Values used for glm on pheromone diversification in sister species. Dissimilarity is expressed by Euclidean distance, and branch length is in million years.

Chapter 2: QTL mapping of pheromone differences in a pair of sympatric *Heliconius* sister species

Abstract

Speciation and natural selection leave observable traces on the genomes of organisms. However, these processes do not affect the genome homogenously: in cases of sympatric speciation-with-gene-flow, genes controlling traits involved in reproductive isolation between the species are expected to be found within areas of greater genome differentiation, in so-called "islands of divergence". For this reason, understanding the genomic architecture of reproductive barrier traits is an important aspect of researching speciation. In several neotropical *Heliconius* butterfly species, it has been demonstrated that colour pattern is the most important factor in reproductive isolation, but more recent work has demonstrated that male androconial sex pheromones also have all the necessary properties to act as reproductive barriers. In this study, I focus on the sympatric sister species *Heliconius elevatus* and *Heliconius pardalinus,* known to have several highly differentiated traits that likely promote reproductive isolation, including differences in androconial secretions. Using captive-bred F2 hybrids genotyped at genome-wide RADseq markers, I identify quantitative trait loci (QTLs) associated with the production of seven androconial compounds that are strongly differentiated between *H elevatus* and *H. pardalinus*; henicosane, tricosane, (Z)-9-henicosene, (Z)-11-icosenylacetate, (Z)-11-icosenylpropionate, (Z)-13-docosenylacetate and hexahydrofarnesylacetone. Coincident QTLs on chromosomes 16, 19 and 20 controlling the production of multiple androconial compounds, and are located near regions of high Fst between the species. The region on chromosome 20 is also known to contain genes controlling androconial secretions in *Heliconius melpomene*. Several candidate genes encoding enzymes involved in fatty acid metabolism, a set of biosynthetic reactions known to produce pheromone components in many Lepidoptera species, are located within the identified genomic intervals.

Introduction

Speciation is the process by which a population diverges into two distinct groups, giving rise to two species (Coyne and Orr, 2004). The definition of species has been the object of much debate over the decades (Mallet, 1995). Classically, the biological species concept (BSC) (Mayr, 1963) defines a species as a reproductively isolated unit; accordingly, conspecifics can interbreed but no heterospecific mating can occur, and if it did, the hybrids would not contribute to gene flow, either by being sterile or by being non-viable. This concept in itself implies complete genome divergence between diverging populations and works under the assumption that the genome, being composed of strongly co-dependent genes, essentially operates as a unit (Mayr, 1963), with speciation affecting large numbers of genes at once and new traits being the result of such co-adaptation. Another implication of the BSC that follows logically from his assumption of divergence across the genome is that species cannot arise as long as any level of gene flow is present between the diverging populations (Mayr, 1963).

However, decades have passed since Mayr defined the BSC, and it has become apparent that it represents an extreme end of a speciation continuum: the spectrum of genetic changes that accumulate as two populations diverge into species, going from a situation of complete gene flow (panmixia), where all individuals can interbreed freely, to becoming fully separate biological species with broadly incompatible genomes (Shaw and Mullen, 2014). In light of this, speciation does not actually need to involve complete reproductive isolation in the first place; Wu defined speciation as the stage where two population would not lose divergence, and will be able to continue diverging, even when coming into contact and maintaining gene flow: this is the genic view of species (Wu, 2001), and is at the basis of ecological speciation.

Under this view, genomes can be considered permeable in the initial and intermediate phases of speciation, and alleles that benefit the two species differently (differential adaptation) will diverge even in the face of gene flow (Wu, 2001; Smadja and Butlin, 2011; Feder, Egan and Nosil, 2012; Supple *et al.*, 2015). This concept of speciation-with-gene-flow, nonetheless, remains controversial due to its counterintuitive nature: gene flow and species divergence are opposing forces, as differential adaptation of genes in the two emergent species, as well as potential associations between loci would be broken down by recombination events brought about by the exchange of genetic material during mating (Smadja and Butlin, 2011; Feder, Egan and Nosil, 2012; Via *et al.*, 2012).

In the genic view of speciation, a relatively small number of "speciation genes" is responsible for upholding separation between two species (Wu and Ting, 2004; Nosil and Schluter, 2011; Feder, Egan and Nosil, 2012; Nosil and Feder, 2012). As technological advancements in the field allowed us to perform genome scans (Feder *et al.*, 2013), we now know that levels of differentiation between

species vary greatly across different regions of the genome (Wu, 2001), a phenomenon known as heterogenous genome divergence (Wu and Ting, 2004; Nosil, Funk and Ortiz-Barrientos, 2009; Nosil and Feder, 2012). Gene flow has a homogenizing effect on genomic regions not involved in differential adaptation, whereas loci that have a key role in reproductive isolation remain unaffected by recombination (Wu, 2001). Natural selection is not the only factor dictating a lack of recombination in certain loci: the very architecture of chromosomes makes exchange of genetic material less frequent in some areas, for example where a rearrangement such as inversions have taken place, or close to the centromeres (Feder *et al.*, 2003, 2013; Slotman *et al.*, 2006; Vincenten *et al.*, 2015).

Recombination-resistant loci, whichever the initial cause of their persistence in the face of gene flow, may be associated with other genes and traits, be it via physical linkage with neighbouring regions of the chromosome or pleiotropic effects, which also diverge alongside them (Smadja and Butlin, 2011; Via *et al.*, 2012). Regions adjacent to selected loci also experience a reduced frequency of crossing over events: thus, divergence at a single speciation locus can also promote divergence of neighbouring regions, a phenomenon known as divergence hitchhiking (Feder *et al.*, 2012; Feder, Egan and Nosil, 2012; Via *et al.*, 2012). Loci with active roles in speciation tend to accumulate within these regions of high differentiation, that have thus been called "islands of divergence" (Feder, Egan and Nosil, 2012). As the species diverge further, these islands grow in size, and selection begins to reduce recombination globally across the genome, producing more and more widespread differentiation in the phenomenon of genome hitchhiking (Feder, Egan and Nosil, 2012). Eventually, the genomes would be incompatible, and gene flow between the species would gradually come to an end.

Speciation-with-gene-flow and genomic islands of divergence are particularly relevant in cases of sympatric speciation, where the two diverging populations have largely overlapping ranges with no geographical barriers to keep them from interbreeding. Sympatric speciation, in the same vein as speciation-with-gene-flow, is a debated topic, however its plausibility is nowadays widely accepted (Dieckmann and Doebeli, 1999; Via, 2001; Bolnick and Fitzpatrick, 2007; Fitzpatrick, Fordyce and Gavrilets, 2008). It has been shown in past studies that species pairs that diverged in sympatry did so without complete interruption to gene flow, a process which leaves measurable traces on the genome: islands of divergence forming around genes under differential selection, and substantial evidence for divergence hitchhiking affecting variable numbers of genomic regions (Feder *et al.*, 2003; Turner, Hahn and Nuzhdin, 2005; Michel *et al.*, 2010; Martin *et al.*, 2013). The genomes of recently diverged, incipient or closely related species, wherein these effects are most noticeable, thus offer an opportunity to investigate the early stages of the speciation continuum, and identifying speciation genes related to ecologically divergent traits and their location relative to islands of divergence is an important part of this endeavour (Mérot *et al.*, 2017).

Heliconius butterflies have become a model system in the study of the speciation continuum (Merrill *et al.*, 2015; Supple *et al.*, 2015; Mérot *et al.*, 2017). Resulting from a recent evolutionary radiation (Kozak *et al.*, 2015), this genus includes many recently diverged sister pairs which likely evolved in different geographic contexts, with different species pairs and intraspecific races potentially representing a variety of stages in the speciation continuum (Rosser *et al.*, 2015; Mérot *et al.*, 2017). In several species such as *Heliconius melpomene* and *Heliconius erato*, there are multiple colour pattern races, distributed parapatrically across the species' ranges. Across hybrid zones between adjacent races, genome divergence between the two races is typically low, consistent with a scenario of gene flow, save for a few islands of divergence (Nadeau *et al.*, 2012; Martin *et al.*, 2013; Van Belleghem *et al.*, 2017). Within these islands, we typically find colour pattern genes (Nadeau *et al.*, 2012; Martin *et al.*, 2013), which control the warning colour patterns that signal the butterflies' unpalatability to predators.

Mate preference in many *Heliconius* species is strongly affected by colour pattern, with individuals showing a strong preference for their own colour pattern (assortative mating) (Jiggins *et al.*, 2001; Naisbit, Jiggins and Mallet, 2001; Salazar *et al.*, 2010; Merrill *et al.*, 2012). This dual role of colour pattern in affecting both mate preference and mimicry means that this trait can be referred to as a "magic trait" that facilitates speciation with gene flow (Servedio *et al.*, 2011). Therefore, shifts in colour pattern are thought to be an important driver of speciation in this group of butterflies.

However, colour pattern preference cannot fully explain reproductive isolation in *Heliconius* (Jiggins *et al.*, 2001). This is particularly evident in cases where sister species belong to the same mimicry, yet show strong assortative mating (Mérot *et al.*, 2015). Therefore, identifying other traits acting as pre-zygotic reproductive barriers (particularly relevant speciation with gene flow) (Coyne and Orr, 1989) has become a point of interest in the study of *Heliconius* and other groups. In *Heliconius,* male sex pheromones are one such candidate trait (Rosser *et al.*, 2019; González-Rojas *et al.*, 2020) (Chapter 1). Male sex pheromones are species-specific (Mann *et al.*, 2017; Darragh *et al.*, 2020), involved in courtship and mate choice (Vanjari *et al.*, 2015; Darragh *et al.*, 2017), able to diversify rapidly and strongly differentiated in sympatric sister species pairs (Chapter 1), all of which make them suitable for a role in reproductive isolation.

In Chapter 1, I demonstrated that sympatric sister pairs have strongly divergent androconial chemical signatures. In this chapter, I explore the genetic architecture of pheromone differences in one sympatric sister pair, *H. elevatus* and *H. pardalinus butleri.* These species have different ecologies, including not only their colour pattern and their pheromones but also their wing shape, flight dynamics (Queste, 2020) and host plant preference (Rosser *et al.*, 2019), yet they occupy overlapping ranges

(Rosser *et al.*, 2015). Save for a few regions where genome divergence (F_{st}) is much higher (Rosser et al., unpublished) (Kryvokhyzha, 2014), the genomes of these two species genomes show very little divergence, a situation consistent with a scenario of gene flow and admixture between close relatives. Within some of these islands or hotspots of divergence, as seen in other species pairs, we find colour pattern genes, which unlike the rest of the genome, do not reflect the two species' taxonomic relationship. Instead, at these genes, *H. elevatus* clusters closer to the more distantly related co-mimetic *H. melpomene,* a result that reveals *H. elevatus'* origin as a possible hybrid species, with its colour patterns having introgressed from *H. melpomene* (Dasmahapatra *et al.*, 2012).

Despite their different ecologies, *H. elevatus* and *H. pardalinus* are able to hybridize and produce fertile F1 offspring, that can cross to produce F2 hybrids, or backcross to either parental species. This permits the investigation the genetics of putative speciation traits as it allows for QTL mapping of species differences. Using F2 hybrids, I investigated whether the very divergent androconia secretions of *H. elevatus* and *H. pardalinus* map to islands of divergence between the two species in the same way as their colour pattern does. Both species mainly produce fatty acid (FA) derived molecules as part of their andrconial blend, with the addition of plant derivatives in the case of *H. pardalinus*. The *H. elevatus* androconial blend has a very simple chemical signature containing one predominant type of FA derivative, alkanes, with its two main products being henicosane and tricosane. *H. pardalinus,* on the other hand, produces a more complex blend where the FA-derived alkene (Z)-9-henicosene and ester (Z)-11-icosenylacetate are most prominent, accompanied by other alkenes and esters as well as plant derivatives, mainly hexahydrofarnesylacetone (Mann *et al.*, 2017)(Chapter 1). My aims were to:

1. Identify QTLs associated with the production of these individual compounds.
2. Search the genomic intervals to identify relevant candidate genes.
3. Test whether the association between genotypes and phenotypes in these compounds reflects the phenotypes of the two pure species
4. Compare my results to previous results from *H. melpomene* (Darragh *et al.*, 2017; Byers *et al.*, 2020, 2021) to look for any evidence of introgression of pheromone genes between *H. melpomene* and *H. elevatus.*

Materials and Methods

Tissue collection

Heliconius elevatus and *Heliconius pardalinus butleri* stocks were established in Tarapoto (Peru) from wild populations collected in San Martin (Peru). While not the only race of its species present in this range, *H. pardalinus butleri* will be referred to simply as *H. pardalinus*. Part of the crosses were carried

out in insectaries in Tarapoto and part in greenhouses at the University of York (York, UK). F1 hybrids were obtained by hand pairing and were then allowed to mate naturally or hand-paired to produce F2 hybrids or backcrosses to both parental species. This work was carried out by and reported in Rosser *et al.* (2019) and Queste (2020).

To collect the relevant tissue samples, butterflies were sacrificed, and from each male two tissue samples were excised from the hindwing using sterile scissors and forceps: the androconia, corresponding to a sliver of tissue at the anterior margin of the hindwing (Emsley, 1963, 1965), and another region of the wing to serve as a negative control. These tissue extracts were suspended in HPLC grade dichloromethane (Sigma-Aldrich, St. Louis, Missouri) alongside 212.8 ng tridecyl acetate to act as the internal standard. The latter was prepared via tridecane esterification in a dichloromethane medium. Immediately following the androconia and control extraction procedure, the hybrids' body was preserved separately in a dimethyl sulfoxide solution (20% DMSO, 0.25 M EDTA, saturated with NaCl) and stored at -20°C until DNA extraction.

Pheromone analysis

Androconia and control tissue extracts were analyzed on a GC-MS system including a 7890A GC-System coupled with an MSD 5975C Mass Analyzer (Agilent Technologies, Santa Clara, USA) fitted with a HP-5MS column (50m length 0.25 mm I.D., 0.25 µm film thickness; Agilent Technologies). The ionization method was electron ionisation with an electron energy of 70 eV. The instrument conditions were as follows: inlet pressure 9.79 psi, He 20 mL min^{-1}, injection volume 1 µL. The GC was programmed with the following temperature ramp: start at 50°C, increased by 5°C min^{-1} to a maximum temperature of 320°C, then held for 5 min making up a total run time of 69 min. All samples were run at the Department of Organic Chemistry at Braunschweig Technische Universität (Braunschweig, Germany). Together with the tissue samples, a C6-C40 alkane ladder was also run for the purpose of retention index (RI) calibration. Data deconvolution and analysis was carried out using the AMDIS (Automated Mass Spectral Deconvolution and Identification System) software (Mallard and Reed, 1997), distributed by NIST (National Institute of Standards and Technology). For all identified compounds, the amount was calculated from the GC peak's area, as reported by AMDIS. To provide a putative identity for each compound, mass spectrometry was interpreted by the AMDIS software via the NIST databases and the additional databases compiled at the Institute of Organic Chemistry of Braunschweig TU by the Schulz lab, also responsible for originally identifying a large number of products present in *Heliconius* androconia (Mann *et al.*, 2017). The RI was used for compound identification and validation alongside the retention time (RT) and the mass spectrometry fragmentation pattern. All compounds between undecane (RI=1100) and octacosanal (detected RI=3041, reported RI=3039.7 (Andriamaharavo, 2014) were scored. This analysis also included F1

hybrids and backcrosses to both *H. elevatus* and *H. pardalinus,* as well as pure species with the full results being reported in (Cama, 2016). However, only F2 individuals were later used for QTL mapping reported here, so all genomic data reported within this chapter refers only to F2 crosses, with past collected data on F1 and backcrosses being included just in discussions pertaining to the phenotypes.

DNA extraction and sequencing

DNA was extracted from a 1/3-to-2/3 section of the F2 hybrid butterflies' thorax depending on the individual's size, using the Qiagen DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany) and following the protocol provided with the kit. From these extractions, restriction site associated (RAD) libraries were prepared using the protocol described in (Baird *et al.*, 2008) with modifications detailed in (Hoffman *et al.*, 2014) mainly by Michaela Nelson (University of York). Sequencing was carried out at FAS Center for Systems Biology (Harvard, Massachusetts) using the Illumina HiSeq 2500 platform (Illumina, Inc., San Diego, California). The resulting sequences were prepared, cleaned and aligned to the *H. melpomene* genome v.2.5 (Davey *et al.*, 2016) by Rosser, N. (University of Harvard) and Dasmahapatra, K. (University of York), and used to assemble a linkage map in Lep-MAP3 (Rastas, 2017), that was then converted into a genotype format and utilized in the subsequent quantitative trait locus (QTL) analysis. The genotypes were expressed as EE (homozygous *H. elevatus* alleles), PP (homozygous *H. pardalinus* alleles) and EP (heterozygous alleles from the two parental species).

QTL analysis

Out of all the phenotyped individuals, genotype data was only collected for 123 from 21 out of the 26 F2 families, and thus only these were used in the analysis. The compound amounts were logged so as to normalize the data, and no further transformation or dimensionality reduction techniques were applied. The QTL analysis was carried out using the R(qtl) package (Broman *et al.*, 2003). Genome scans for single-locus QTLs were carried out for each individual compound using the scanone function, using the normal phenotype model and the Haley-Knott (HK) regression method (Haley and Knott, 1992). This function outputs a LOD score for each tested compound phenotype at each marker position in the genome, a value that signifies the likelihood of association between a phenotype and a marker, with high LOD scores indicating the likely location for the QTLs controlling that compound. The 95% confidence intervals for each QTL peak located in this manner were calculated using the bayesint function on the output of scanone. The 95% and 99% significance thresholds for each detected QTL were determined using the same scanone model as for the single-locus QTL scans, but set up to perform 1000 permutations.

I calculated the proportion in phenotypic variance explained by each QTL using the formula $1 - 10^{-2\,LOD\,/\,n}$ where n is the sample size and LOD is the LOD score output from scanone, as described in (Broman *et al.*, 2003). The proportion in phenotypic variance explained by a QTL is a component of heritability, which describes how much of a trait's variance can be explained by genetic inheritance (Tang *et al.*, 2018). Values close to or equal to 1 imply a strong contribution of the QTL to the trait's phenotype, whereas values closer to 0 imply a weak contribution. The latter scenario does not necessarily mean that the trait is not heritable, as trait heritability is given by the combined contribution to phenotypic variance of all QTLs associated with the trait rather than individual ones. Thus, when the proportion of variance explained is small, a more likely interpretation is that the trait's phenotype is produced by multiple QTLs, some of which may not have been detected.

The concentration of compounds in pure *H. elevatus* and *H. pardalinus* is known (see Chapter 1), so in order to test whether hybrid phenotypes are associated with the correct genotypes, phenotype and genotype information for each compound with significant QTLs were extracted using pull.pheno and pull.genoprob respectively. Under a simple additive model, F2 individuals with an EE genotype are expected to have *H. elevatus*-like phenotypes, PP individuals are expected to have *H. pardalinus*-like phenotypes, and EP individuals are expected to have F1 phenotypes. Lastly, the genotype:phenotype effect for each significant QTL was plotted using plotPXG.

To test whether the detected QTLs fall within regions of high genome divergence between the species, their location was plotted onto data for genomic differentiation data calculated as sliding window F_{st} across the genome (provided by Rosser, N.). Having identified the location and confidence intervals for each significant QTL, I located and extracted the corresponding regions from the *Heliconius melpomene* genome Hmel2.5 on the Lepbase database (Challi *et al.*, 2016) and used protein BLAST (BLASTp) to characterize the genes found within those regions against the SwissProt protein database, looking for enzymes relevant to *Heliconius* pheromone biosynthesis as described in (Mann *et al.*, 2017), mainly those potentially involved in fatty acid (FA) metabolism (Reed *et al.*, 1994; Tillman *et al.*, 1999; Liénard *et al.*, 2014; Mann *et al.*, 2017; Byers *et al.*, 2020). For a breakdown of the biosynthetic reactions involved in the production of these compounds, and the relevant enzymes expected to find within the confidence intervals, see Supp. Fig. 2.1.

Results

Hybrid extracts phenotypes

I collected phenotype data from 151 male F2 hybrids from 26 different families between those bred in Tarapoto (Peru) and York (UK). In total, the number of different chemical compounds detected in the dataset, excluding likely contaminants, was 40. Of these, 25 compounds show significant

differences between the two parental species *H. elevatus* and *H. pardalinus,* as well as being significantly different between male androconia and controls (Rosser *et al.*, 2019), and are therefore good candidates as pheromone components. Since these compounds have not been tested for any behavioural effects, I refer to these compounds as androconial compounds and androconial blends rather than pheromones. Several types of compounds are present in this roster: alkanes, alkenes and esters feature prominently, with the first being the main *H. elevatus* product and the latter two being typical of *H. pardalinus* (Cama, 2016; Mann *et al.*, 2017). *Heliconius pardalinus* products also include the plant derivatives phytol and its degradation product hexahydrofarnesylacetone (HHFA), and both species, produce the plant derivatives homovanillylalcohol and syringaldehyde, though at different levels, with only the former being significantly different between the two. See Supp. Table 2.1 for a list of all 25 compounds included in this study with their respective mean amounds.

F2 hybrids typically contain products from both parental species rather than showing decidedly *H. elevatus* or *H. pardalinus*-like chemical profiles, however the distributions of amounts these products take on among the F2 individuals varies greatly. Some, like the alkanes henicosane and (Z)-11-icosenylacetate, show signs of incomplete dominance of the parental alleles, taking on a range of values that vary between the typical *H. elevatus* and *H. pardalinus* amounts (Fig. 2.3). This is further validated by the F1 phenotype, that shows an intermediate value between the two pure species, and the backcrosses, which show values close to their respective pure species parent. Others, like the alkenes (Z)-9-henicosene and (Z)-9-tricosene, show signs of dominance for one of the parental alleles, *H. pardalinus* in this example, with F1 hybrids having phenotypes most similar to *H. pardalinus,* but actually surpassing the amounts observed in the pure species. Others still, like homovanillylalcohol, show greatly enhanced levels of production that greatly surpass both pure species, both in the F2 and in the F1, with backcrosses also showing distinct phenotypes from their parental pure species (Supp. Table 2.1).

These findings are accurately summed up when plotted as a PCA to get a general view of the variety of phenotypes observed in the F2s compared to pure species, backcrosses and F1s (Supp. Fig. 2.2). The PCA clearly shows that F2 individuals, particularly those that originated from Peru families, display an extremely high variety in phenotypes compared to the parental species, the backcrosses and the F1s, with the PC2 axis explaining mainly variation within the F2. This variety in chemical profiles hint at potential difficulties with QTL mapping, as compound phenotypes that are not typical of either parental species would probably result from complex interactions between loci and thus some compounds may not accurately be mapped to a single region of the genome.

QTL mapping of hybrid chemical profiles

Having identified 25 compounds that showed patterns consistent with putative pheromone substances (significantly different between species, between sexes, and between androconia and control), genome scans were carried out on every individual compound. 123 fully genotyped and phenotyped F2 individuals were used for this purpose. *H. elevatus* and *H. pardalinus* F2 crosses sport a wide variety of chemical profiles, and most of the compounds could not be mapped to any specific locus of the genome. Significant QTLs (significant at least at the 0.05 level) were detected for 7 of the 25 compounds that are significantly different between the species (Cama, 2016; Rosser *et al.*, 2019) (Fig. 2.1). These compounds include henicosane and tricosane (Fig. 2.3A-B), the two main *H. elevatus* products, and the two most prominent *H. pardalinus* products (Z)-9-henicosene and (Z)-11-icosenylacetate (Fig. 2.3C-D), alongside three other typical *H. pardalinus* compounds that are almost never observed in *H. elevatus* (Fig. 2.3E-G), (Z)-11-icosenylpropionate, (Z)-13-docosenylacetate and hexahydrofarnesylacetone. Of these, henicosane, (Z)-9-henicosene and (Z)-13-docosenylacetate were also significant at the 0.01 level. Uniquely among these 7 compounds, hexahydrofarnesylacetone showed two peaks, on chromosome 19 and on chromosome 20, albeit the latter was only significant at the 0.1 level. No significant QTLs were detected for the remaining 18 compounds, and this is partly a consequence of some of these compounds not being strongly differentiated between the pure species, and in some cases are not even present in all pure individuals of *H. elevatus* or *H. pardalinus*.

The QTL peaks vary in size, location and significance level (Fig. 2.1). Most of them are located on chromosome 19, occupying a large area between the genomic markers Hmel219001o:3680 and Hmel219002o:141794. Within this extensive region we find significant QTLs for henicosane (Hmel219001o:12591008-219002o:141794), hexahydrofarnesylacetone (Hmel219001o:3373019-Hmel219002o:141794) and the overlapping (Z)-11-icosenylacetate (Hmel219001o:3680-10514853) and its direct derivative (Z)-11-icosenylpropionate (Hmel219001o:3680-Hmel219001o:4848438). A single compound, (Z)-9-henicosene, maps to a QTL on chromosome 20 between markers Hmel220003o:11832016 and Hmel220003o:14584000. The Hmel220003o linkage group is a known location for several QTLs that describe pheromone differences between *H. melpomene* and *H. cydno* (Byers *et al.*, 2021). The last two remaining compounds, tricosane and (Z)-13-docosenylacetate, map to chromosome 16, between markers Hmel216002o:5395909-10031338 and Hmel216002o:10021318-10031614 respectively. A full summary of these results is reported in Table 2.1.

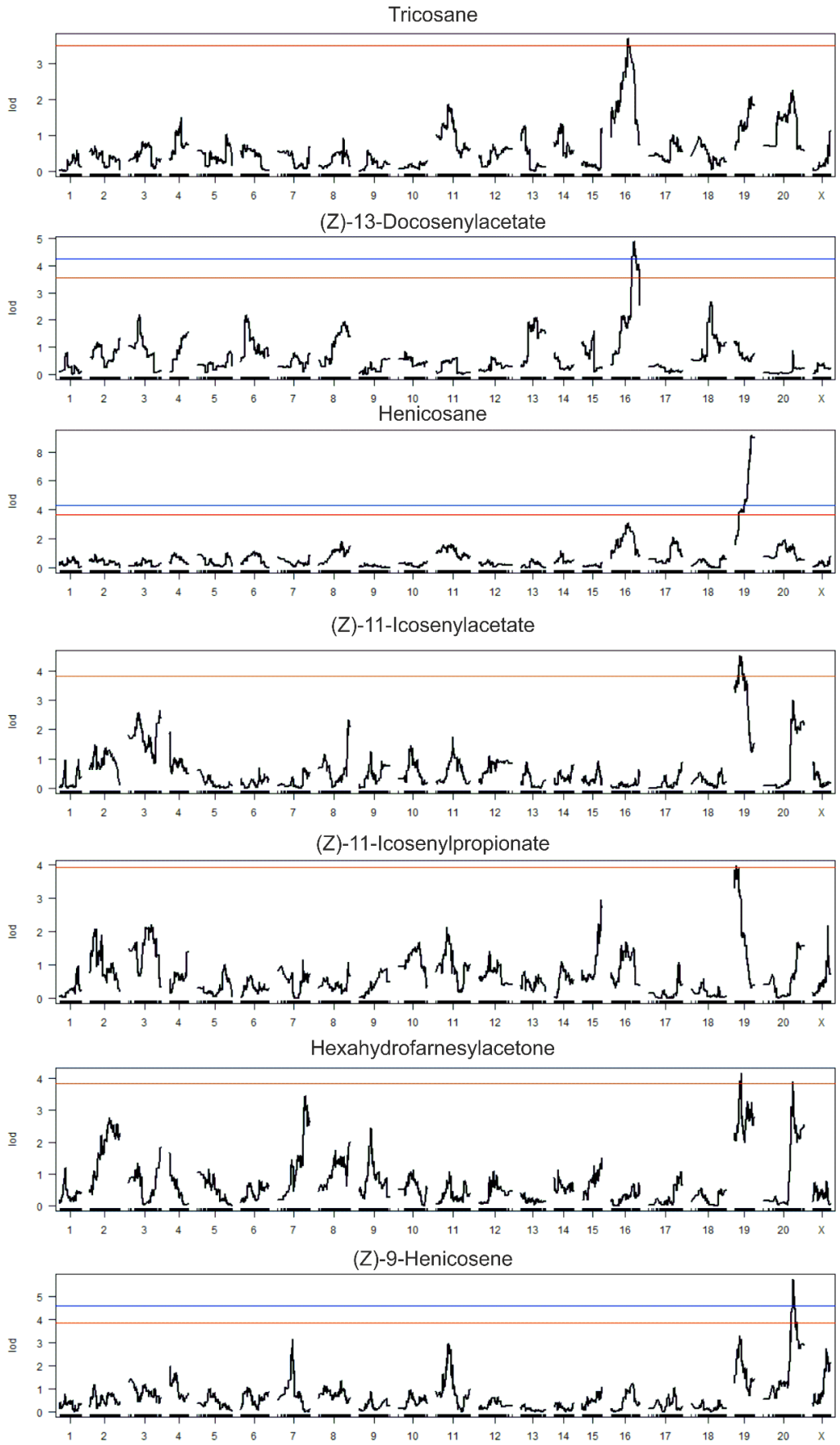


Figure 2.1. All significant QTLs associated with major *H. elevatus* and *H. pardalinus* compounds. Horizontal lines indicate significance levels (0.05 for the red line, 0.01 for the blue line).

Candidate pheromone synthesis genes

As the confidence intervals for the significant QTLs can be very large, the number of protein coding genes found within them is often extremely high. However, for most of the compounds with significant QTL peaks, BLASTp identified several suitable candidate genes. The chromosome 19 region henicosane maps to contains 10 fatty acyl-CoA Δ9-desaturases from the house cricket *Acheta domesticus,* while the (non overlapping) region that contains (Z)-11-icosenylacetate includes 8 putative fatty acyl-CoA reductases from *Drosophila melanogaster* as well as a fatty acyl-CoA Δ11-desaturase from the moths *Trichoplusia ni* and *Spodoptera littoralis*; however, while the (Z)-11-icosenylacetate region fully encompasses the comparatively smaller (Z)-11-icosenylpropionate confidence interval, none of these relevant genes were found within the latter. The hexahydrofarnesylacetone QTL spans an extremely large confidence interval that spans those of both (Z)-11-icosenylacetate and henicosane, however no genes directly relevant to phytol metabolism were found within the region. All of the confidence intervals for significant QTL peaks on chromosome 19 also contain several cytochrome P450 enzymes (10 in the henicosane confidence interval, 8 in the (Z)-11-icosenylacetate one) from *Drosophila,* the moth *Helicoverpa zea,* the butterflies *Papilio glaucus* and *P. polyxenes,* and the cockroaches *Blattella germanica* and *Blaberus discoidalis*. These enzymes mediate important biosynthetic steps in the conversion of precursor FAs to their pheromone derivatives (Tillman *et al.*, 1999).

On the other hand, the chromosome 20 region (Z)-9-henicosene maps to contains 5 putative fatty acyl-CoA reductases from *D. melanogaster,* 3 of which (corresponding to transcripts HMEL035367g1.t1, HMEL035374g1.t1 and HMEL035375g1.t1) had been previously reported within the confidence interval for the main *H. melpomene* compound 1-octadecanol (Byers *et al.*, 2020). Unlike chromosome 19, the (Z)-9-henicosene interval on chromosome 20 carries no genes related to cytochrome P450. Compared to these chromosome 19-20 clusters of suitable gene candidates for pheromone biosynthesis, the chromosome 16 region tricosane and (Z)-13-docosenylacetate map to appears largely devoid of relevant genes, with the only fatty acid metabolism-related gene being a long chain acyl-CoA synthetase from *Arabidopsis thaliana.*

When plotted against *H. elevatus-H. pardalinus* genome divergence (Fst) (Fig. 2.2), some of these significant QTLs fall within or close to areas of high genomic divergence between the two species. With the exception of (Z)-13-docosenylacetate, which maps to an extremely small interval in the telomeric region of chromosome 16, all other compounds' confidence intervals include at least one region of divergence.

Phenotype-genotype associations for each compound usually behaved in an intuitive manner, with EE F2 individuals showing more *H. elevatus*-like concentration, PP more *H. pardalinus*-like and EP taking

on intermediate values (Fig. 2.3). The exceptions to this are tricosane (Fig. 2.3B), that shows a higher predicted concentration in PP individuals whereas it is expected to be present at higher amounts in EE individuals, and (Z)-13-docosenylacetate (Fig. 2.3F), a *H. pardalinus* compound that shows the opposite pattern, albeit unlike tricosane, this odd result may be given by poor data: (Z)-13-docosenylacetate is usually absent in the F2 and can be absent in the parental *H. pardalinus.*

The proportion of variance explained by each QTL, shown in Figure 3, was relatively low for all compounds that gave significant peaks. Out of all compounds, the QTL for henicosane, which produced the clearest peak, also explains the largest proportion of variance in its respective trait, with a value of 0.29. All the others produce values ranging from 0.13 (tricosane) to 0.19 ((Z)-9-henicosene). This may indicate that genetics are a relatively poor predictor of a compound's phenotype, the latter depending more on environmental factors, but the most likely explanation is that each compound depends on multiple QTLs and therefore, the proportion of phenotypic variance explained by a single QTL does not fully reflect the heritability of the trait.

Discussion

Investigating the genetic architecture of traits underlying reproductive isolation is fundamental to our understanding of speciation. *Heliconius* sympatric sister species pairs have been previously shown to produce vastly divergent chemical blends from their androconia (Chapter 1), a major point in favour to the argument that male sex pheromones emitted during courtship may act as reproductive barriers. Here, I investigated the genetics of pheromone composition in two sympatric sister species, *H. elevatus* and *H. pardalinus.* Male sex pheromones are just one of the several strongly divergent traits involved in pre-zygotic isolation between these two species (Rosser *et al.*, 2019; Queste, 2020). The genomes of the two species show overall very low divergence, consistent with a scenario of pervasive ongoing admixture, yet several divergence hotspots with high F_{st} values can be observed (Kryvokhyzha, 2014). Past research has shown that the colour pattern loci *N/Yb* and *B/D* fall within two such islands of divergence on chromosome 15 and 18 respectively (Dasmahapatra *et al.*, 2012), both of which were produced by introgression of colour pattern loci from *H. melpomene* to *H. elevatus* (Dasmahapatra *et al.*, 2012). The resulting divergent colour patterns are a reproductive barrier between the two species as each species displays assortative mating behaviours (Rosser *et al.*, 2019).

Many islands of divergence between the *H. elevatus* and *H. pardalinus* genomes, however, remain unaccounted for. In this study, I located the QTLs for pheromone differences between the two species, using F2 hybrid genomes, and I was able to demonstrate that two such islands contain loci potentially linked to putative pheromone components production. I found a large number of relevant genes involved in FA metabolism (Supp. Fig. 2.1A) within two main divergence islands on chromosome 19

and 20, and the fact that these traits map to areas of high genomic differentiation (Fig. 2.2) is another argument in favour of their role in reproductive isolation between these two sympatric sister species.

Pheromone QTLs map to divergence hotspots containing relevant genes

The androconial blends of *H. elevatus* and *H. pardalinus* are extremely divergent. Both species produce fatty acid (FA) derivatives as part of their pheromones (Mann *et al.*, 2017), but *H. elevatus* mainly produces alkanes whereas *H. pardalinus* produces mainly alkenes and esters, and an overall greater variety of compounds (Mann *et al.*, 2017) including plant derivatives such as phytol and its degradation product (Schulz, Yildizhan and van Loon, 2011) hexahydrofarnesylacetone (Chapter 1). In total, of the 25 compounds detected in the androconial blends of the two species that are likely pheromone candidates (Rosser *et al.*, 2019), I found 7 compounds to have significant QTL peaks, across three chromosomes: 16, 19 and 20.

Chromosome 19. Many FA metabolism-related genes can located in the genomic intervals identified on chromosome 19. The main *H. elevatus* product, the straight-chain alkane henicosane, maps to chromosome 19 alongside three *H. pardalinus* products: the esters (Z)-11-icosenylacetate and (Z)-11-icosenylpropionate, and the plant-derived hexahydrofarnesylacetone. Henicosane's confidence interval encompasses a particularly dense divergence hotspot which includes a large number of potentially pheromone-related genes, the most striking of which are five insect fatty acyl-CoA Δ9-desaturases. Δ9-desaturases are enzymes responsible for introducing a double bond into a fatty acyl-CoA precursor (Supp Fig. 2.1A), likely utilized by *Heliconius* to produce its desaturated pheromone components (Mann *et al.*, 2017), making them directly connected to the biosynthesis of (Z)-9-henicosene, despite the latter mapping to chromosome 20 rather than chromosome 19. It is possible that henicosane maps to a region dense in Δ9-desaturase loci because its production may be linked to a deactivation of these genes, meaning a lack of or a reduction in Δ9-desaturase activity leads to henicosane biosynthesis. This hypothesis seems corroborated by the relative levels of henicosane and (Z)-9-henicosene in the pure species: both species are capable of producing both compounds, but *H. elevatus* produces high levels of henicosane and low levels of (Z)-9-henicosene, while the reverse is true for *H. pardalinus* (Cama, 2016; Mann *et al.*, 2017) (Chapter 1), so the two compounds representing alternative end products to the same pathway seems like a likely explanation.

The esters (Z)-11-icosenylacetate and (Z)-11-icosenylpropionate map to a different region of chromosome 19 that also includes a fairly dense divergence hotspot (Kryvokhyzha, 2014) (Fig. 2.2). The most notable genes in this confidence interval are 6 fatty acyl-CoA reductases, that mediate the transformation of precursor FAs into fatty alcohols (Liénard *et al.*, 2014; Mann *et al.*, 2017), which can then easily be converted into the final ester products (Bjostad and Roelofs, 1981; Tillman *et al.*, 1999).

It is worth noting however that none of these FA reductases are among those attributed to moth pheromone gland FA metabolism as listed in (Liénard *et al.*, 2014). A moth fatty acyl-CoA Δ11-desaturase is also found within this region. Previous studies proposed that the double bond is introduced in the FA precursor uniquely via Δ9-desaturases (Tillman *et al.*, 1999; Liénard *et al.*, 2014), and the product is then elongated (Mann *et al.*, 2017) (Supp. Fig. 2.1A), however Δ11-desaturases are common in Lepidoptera (Bjostad and Roelofs, 1981; Tillman *et al.*, 1999; Rodríguez *et al.*, 2004; Wang *et al.*, 2014; Groot *et al.*, 2019) and the position of the double bond in (Z)-11-icosenylacetate and (Z)-11-icosenylpropionate suggests that this kind of desaturase might be active in *Heliconius* alongside Δ9-desaturases.

Throughout these chromosome 19 confidence intervals I found large numbers of cytochrome P450 genes, also important in FA metabolism. This large family of proteins mediates oxidation reactions on aldehydes, transforming them into hydrocarbons (alkenes or alkanes depending on whether desaturation has occurred upstream) by removing their carbonyl group (Blomquist, Dillwith and Adams, 1987; Reed *et al.*, 1994, 1995; Tillman *et al.*, 1999).

This study represents the first instance of *Heliconius* sister species pheromone differences mapping to chromosome 19, and it seems to be an important hotspot for many FA metabolism genes. This finding is consistent with past work on QTL mapping of *H. melpomene/H. cydno* pheromone differences: while no *H. melpomene/H. cydno* compound maps to chromosome 19, Byers et al. (2021) reported dense clusters of FA metabolism-related genes on this chromosome. Being able to link these clusters to the production of *H. elevatus/H. pardalinus* androconial blend components is both novel and fascinating.

Chromosome 20. The *H. pardalinus* compound (Z)-9-henicosene maps to a region on chromosome 20 where a divergence island is located. Within this interval, much like in chromosome 19, I found 5 *Drosophila* putative fatty acyl-CoA reductases that may mediate the transformation of FA precursors into aldehyde intermediates which can then be further converted into other products (Liénard *et al.*, 2014; Mann *et al.*, 2017). Unlike chromosome 19, however, the confidence interval found on chromosome 20 contains no cytochrome P450 genes. Chromosome 20 has been identified in past studies as the location of the QTLs linked to major *H. melpomene* pheromone components 1-octadecanol and 1-octadecanal (Byers *et al.*, 2020), as well as several other compounds that show significant differences between the sister species *H. melpomene* and *H. cydno*, and carries a cluster of pheromone biosynthesis-related genes (Byers *et al.*, 2020, 2021). *H. melpomene* lacks (Z)-9-henicosene, however a large number of its androconial compounds map to chromosome 20, including henicosane (Byers *et al.*, 2021), which maps to chromosome 19 in our species. The confidence interval for (Z)-9-henicosene in my study only partially overlaps with that of *H. melpomene,* yet 3 of the 5

reductases found within it are in common were also reported in (Byers *et al.*, 2020, 2021). Interestingly, chromosome 20 also carries the QTL of another trait that shows strong divergence between *H. elevatus* and *H. pardalinus,* the wing shape (Queste, 2020), implying potential genetic linkage between the two.

Chromosome 16. The *H. elevatus* compound tricosane and the *H. pardalinus* compound (Z)-13-docosenylacetate map to chromosome 16. (Z)-13-docosenylacetate produces a very small confidence interval that does not overlap with any divergence hotspots and spans only 18 exons, none of which appears related in any way to FA metabolism. Tricosane spans a much wider area but only includes one FA metabolism-related gene, a peroxisomal long chain acyl-CoA synthase (LACS7 from *Arabidopsis thaliana*) which transforms precursor FAs into their fatty acyl-CoA form (Fulda *et al.*, 2002; Shockey, Fulda and Browse, 2002) to then be further transformed into downstream pheromone products (Liénard *et al.*, 2014; He *et al.*, 2017; MacLean *et al.*, 2018). This gene may be relevant as phytol needs to be converted to phytanoyl-CoA form before further degradation into hexahydrofarnesylacetone (Van Den Brink *et al.*, 2005; Schulz, Yildizhan and van Loon, 2011) (Supp. Fig. 2.1). There is a narrow divergence hotspot within the tricosane confidence interval, though the largest island of divergence on chromosome 16 is found outside of it (Kryvokhyzha, 2014).

Hybrid phenotype-genotype associations

F2 hybrids sport a great variety of androconial blend compositions, with most individuals producing compounds from both parental species as opposed to having a distinctly *H. elevatus* or *H. pardalinus*-like chemical signature (Cama, 2016). Taken singly however, each compound has a diverse range of phenotypes, which can individually be more *H. elevatus*-like or more *H. pardalinus*-like, with a full spectrum of intermediate situations. Knowing this, I tested whether the *H. elevatus* genotype EE and the *H. pardalinus* genotype PP, as well as the heterozygous EP, are associated with the expected phenotypes in the seven compounds for which I detected significant QTL peaks. It must be noted however that barring the henicosane QTL, which explained a relatively large proportion of phenotypic variance (0.29), the six other QTLs explained much smaller proportions of variance. This is likely due to a contribution to phenotypic variance from undetected QTLs, implies that the patterns of inheritance of a compound's phenotype may not be fully predictable based on the proportion of phenotypic variance explained by a single QTL, as these traits are probably produced by multiple loci.

Significant genotype:phenotype associations were found for all seven compounds of interest, and are summed up in (Fig. 2.3). Henicosane and tricosane, the two main *H. elevatus* alkanes, show very similar patterns: their amounts are high in *H. elevatus* and its backcrosses, low in *H. pardalinus* and its backcrosses, and take on intermediate values in the F1 individuals. F2 individuals that carry the EE

genotype at the henicosane QTL produce high amounts of this compound, PP individuals produce low amounts and EP individuals produce intermediate amounts, a result that is in accordance to what we know about the pure species phenotypes. Tricosane on the other hand gives discordant results, with the higher amounts being found in PP individuals rather than EE individuals, a wholly unexpected pattern, especially as tricosane and henicosane are strongly correlated to each other in F2 hybrids (Cama, 2016).

The 5 other compounds with significant QTLs were *H. pardalinus* products: (Z)-9-henicosene, (Z)-11-icosenylacetate, (Z)-11-icosenylpropionate, (Z)-13-docosenylacetate and hexahydrofarnesylacetone. All of these compounds except for (Z)-13-docosenylacetate showed expected genotype:phenotype patterns, with PP individuals producing significantly higher amounts of the compounds compared to EE individuals. (Z)-13-docosenylacetate, on the other hand, in the same manner as tricosane, shows an "inverted" relationship where PP individuals produce lower amounts than EE individuals.

The genotype:phenotype relationship of tricosane and (Z)-13-docosenylacetate is unexpected and difficult to explain. The most likely explanation for the discrepancy is that these two compounds may be produced by interactions between more than one QTL, a scenario that the scanone function R(qtl) would not be able to detect since it is not designed to search for interactions (Broman *et al.*, 2003). If the compounds are produced by more than a single locus, their significant peaks would not be fully representative of their genomic architecture, only to part of it. Both QTLs fall on chromosome 16, in a region of low species divergence and much poorer in putative pheromone-related loci compared to the regions found on chromosome 19 and 20, so it is also possible that these QTLs may be false positives resulting from multiple testing. Additionally, the low proportion of phenotypic variance explained by the detected QTLs forthese traits (0.13 for tricosane and 0.17 for (Z)-13-docosenylacetate) (Fig. 2.3B, 3F) strongly implies that these compounds' production may depend on multiple loci, including undetected ones, which seems like a realistic scenario given the several biosynthetic processes involved in their production. Nonetheless theother 5 significant QTLs explained small proportions of phenotypic variance as well and did not produce discordant genotype:phenotype associations.

Exploring the potential for introgression from *H. melpomene*

H. elevatus acquired its colour pattern via hybridization with *H. melpomene,* through introgression of genomic regions surrounding the *B/D* and *Yb* colour pattern loci on chromosomes 18 and 15 (Dasmahapatra *et al.*, 2012). There is a tangible benefit to horizontal gene transfer of colour pattern alles: *Heliconius* colour patterns are aposematic and different species use Müllerian mimicry to share signalling costs to predators, so introgression can facilitate rapid convergence (Dasmahapatra *et al.*,

2012; Zhang *et al.*, 2016). The same benefit would not apply to pheromones, whose composition should ideally be species-specific (Darragh *et al.*, 2020), as *Heliconius* are not known to benefit from any form of chemical mimicry. Nonetheless, if we assume pheromones act as reproductive barriers, especially given their strongly divergent composition among sympatric sister species pairs (Chapter 1), introgression of pheromone loci between distantly related species can still prove beneficial in rapidly producing large changes in pheromone composition that facilitate differentiation in sympatry. Moreover, pheromone loci may introgress simply by being linked to colour pattern loci, and in fact the idea of associations between mimicry ring and pheromone composition has been proposed and tested in some species in the past (Mann *et al.*, 2017; Ehlers *et al.*, 2021).

However, unlike their colour patterns, the *H. elevatus* and *H. melpomene* pheromone blends are not alike. *H. elevatus* contains mostly alkanes, predominantly henicosane and, in much smaller amounts, tricosane (Rosser *et al.*, 2019)(Chapter 1), whereas *H. melpomene* produces mainly alcohols and aldehydes, especially the alcohol 1-octadecanol and the aldehydes octadecanal, (Z)-11-icosenal and (Z)-13-docosenal (Vanjari *et al.*, 2015; Mann *et al.*, 2017; Byers *et al.*, 2020, 2021; Darragh *et al.*, 2020), and henicosane is actually a predominant compound of *H. cydno* rather than *H. melpomene* (Mann *et al.*, 2017; Byers *et al.*, 2021)(Chapter 1). The main *H. melpomene* product 1-octadecanol is occasionally seen in *H. elevatus,* in *H. elevatus* backcrosses and in F2 hybrids, but the other compounds are never seen in *H. elevatus,* and (Z)-11-icosenal occasionally appears in *H. pardalinus* instead, though it does appear in much larger amounts in some F2 hybrids.

However, the difference in composition between *H. elevatus* and *H. melpomene* does not bar the possibility pheromone genes may have introgressed between the two. All of these compounds derive from the same shared fatty acid metabolic pathway (Tillman *et al.*, 1999; Liénard *et al.*, 2014; Mann *et al.*, 2017). So while the end products of pheromone biosynthesis are different for each species, an introgression event that provided *H. elevatus* with the chromosome 20 reductases from *H. melpomene* may potentially generate phenotypes that are difficult to predict given the complexity of the FA metabolic pathway. Indeed, there are some introgressed regions throughout the divergence islands of chromosome 20, some within the confidence interval I found for for (Z)-9-henicosene (Rosser, Pers. Comm.). There are also numerous introgressed regions on chromosome 19 within the confidence intervals I identified (Rosser, Pers. Comm), however the only *H. melpomene*/*H. cydno* androconial blend compound that maps to this chromosome is benzyl cyanide, not found in *H. elevatus*. Overall, it is not possible to rule out the introgression of pheromone alleles from *H. melpomene* to *H. elevatus* as causing *H. elevatus* to have a divergent pheromone blend to that of its closely-related sister species. *H. pardalinus*.

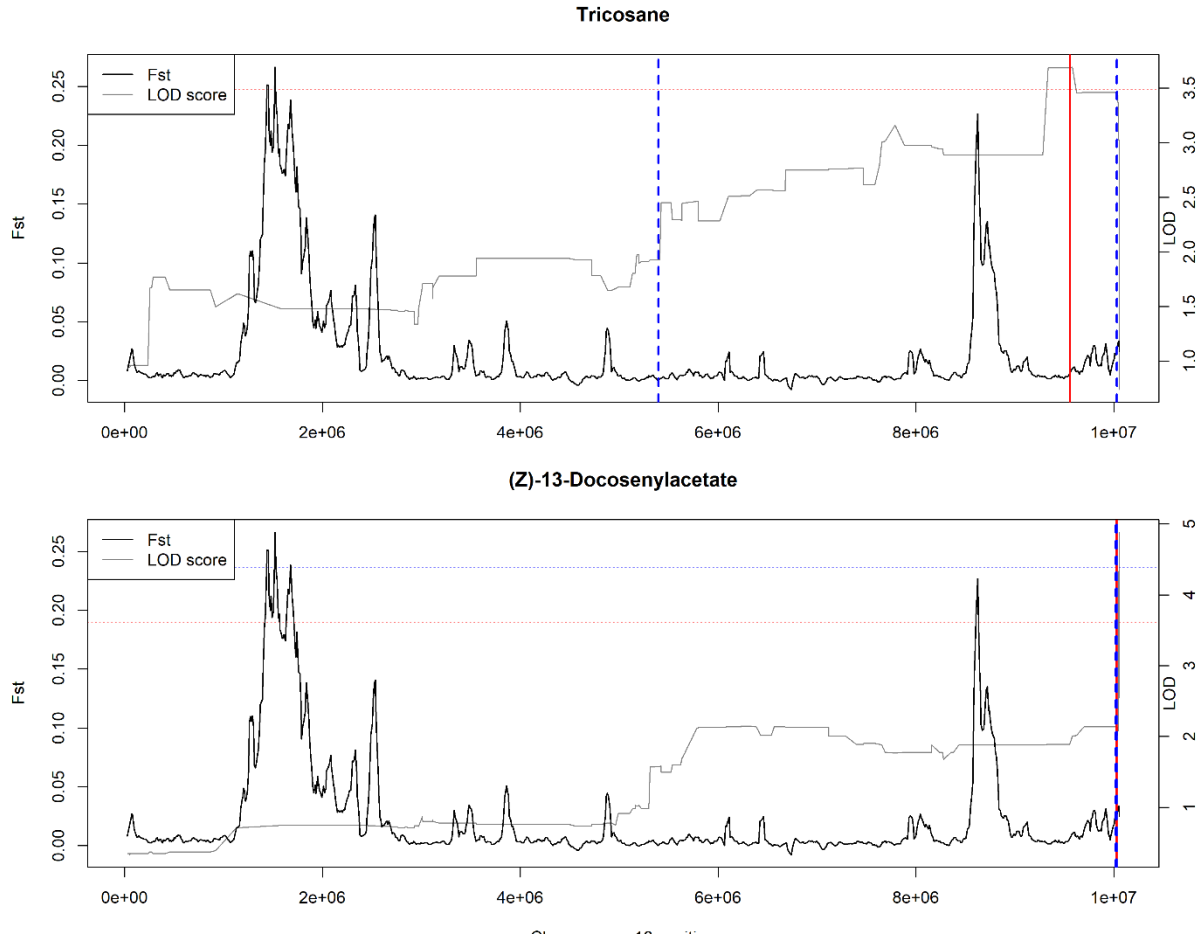
Conclusion

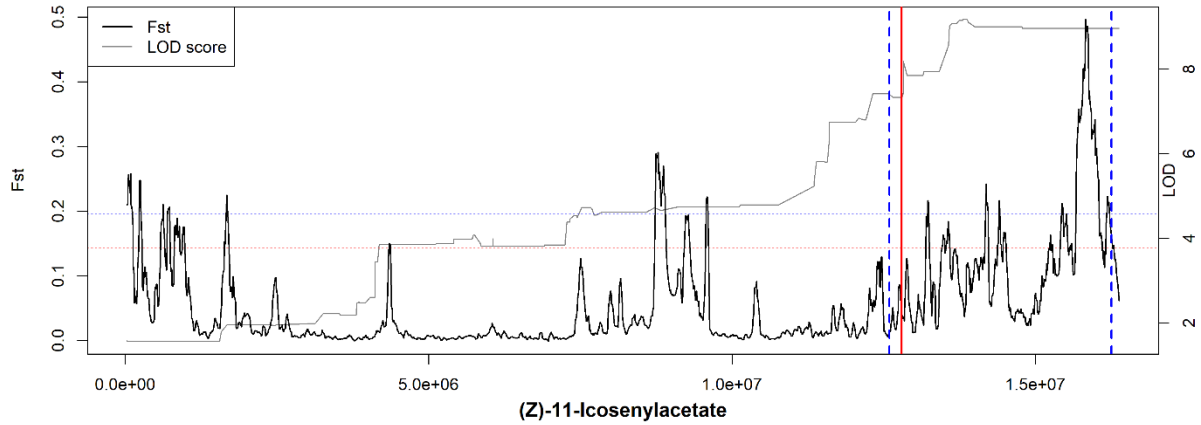
Work on the genetics of *H. elevatus* and *H. pardalinus* pheromones is far from complete. While I was able to find QTLs for the main products in these two species, many androconial blend components (and on the other hand, many enzymes involved in the fatty acid metabolic pathway that leads to pheromone production) remain unaccounted for. Furthermore, my findings seem to suggest that some of the traits may result from interactions with multiple QTLs, something which I was not able to ascertain as of the time of writing.

In spite of these missing answers, and despite the complexity and great variability of androconial blend phenotypes in F2 hybrids as well as in pure species, my results are promising and informative as to the genetic architecture of these traits. Despite the relatively small sample size, I successfully identified several candidate QTLs for the main compounds making up pheromone differences in the two sympatric sister species *H. elevatus* and *H. pardalinus.* Within the confidence intervals I found on chromosome 19 and 20, I identified multiple genes with a likely role in pheromone production, including fatty acid Δ9-desaturases, fatty acid reductases and cytochrome P450 enzymes. The confidence intervals largely (but not always) overlap with hotspots of genomic divergence between the two species, consistent with the putative role of pheromones as a prezygotic reproductive barrier (Rosser *et al.*, 2019)(Chapter 1). One of these islands of divergence, on chromosome 20, also contains QTLs that control wing shape (Queste, 2020), and was previously reported as the location for *H. melpomene/H. cydno* pheromone QTLs (Byers *et al.*, 2020, 2021). However, while the *H. elevatus* colour pattern genes were acquired via hybridization with *H. melpomene,* I found no conclusive evidence for introgression of pheromone loci between these species.

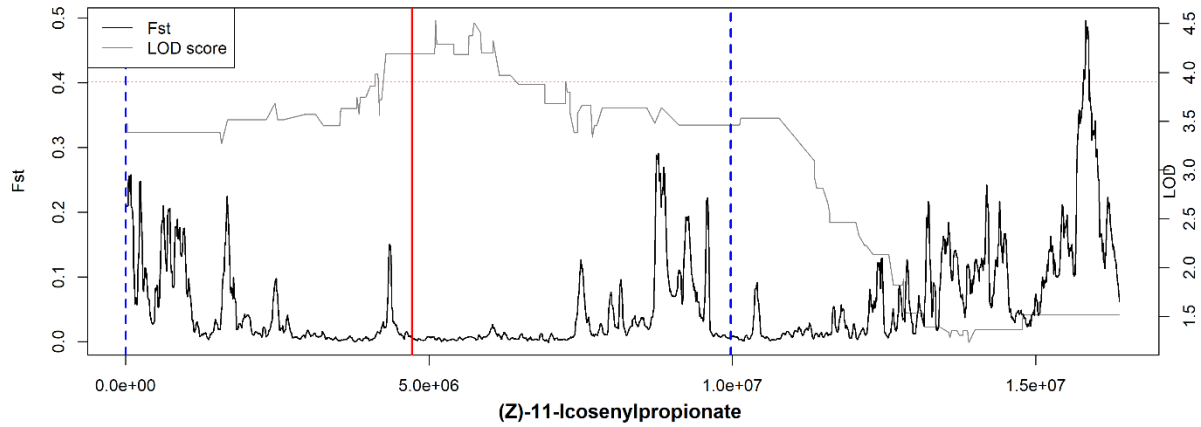
Figures and Tables

Figure 2.2. 95% confidence intervals for compounds that map to QTLs on chromosome 16, 19 and 20 with CI boundaries represented by dashed vertical lines, overlaid onto plots of genome differentiation between *H. elevatus* and *H. pardalinus,* expressed by the Fst value (black line) representing genome divergence. The LOD score for each QTL is shown as a grey line.

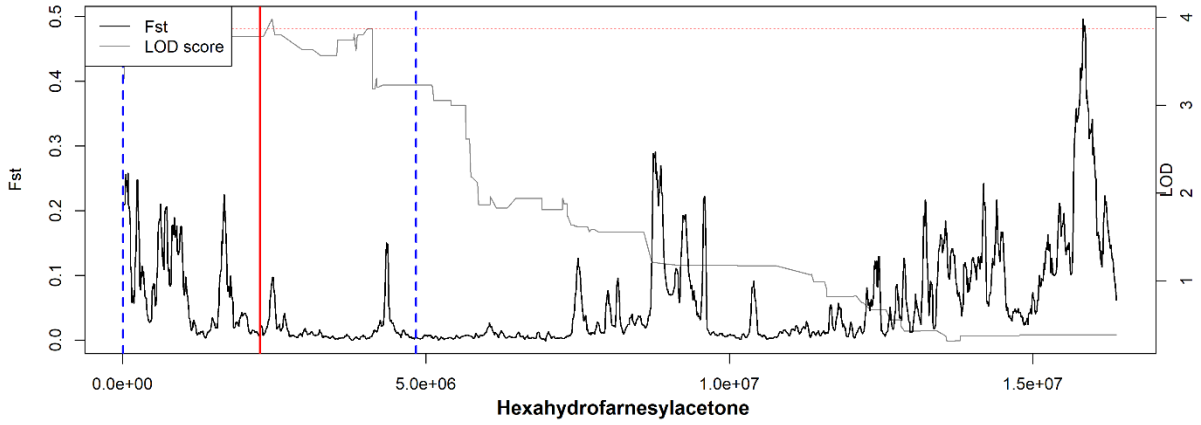




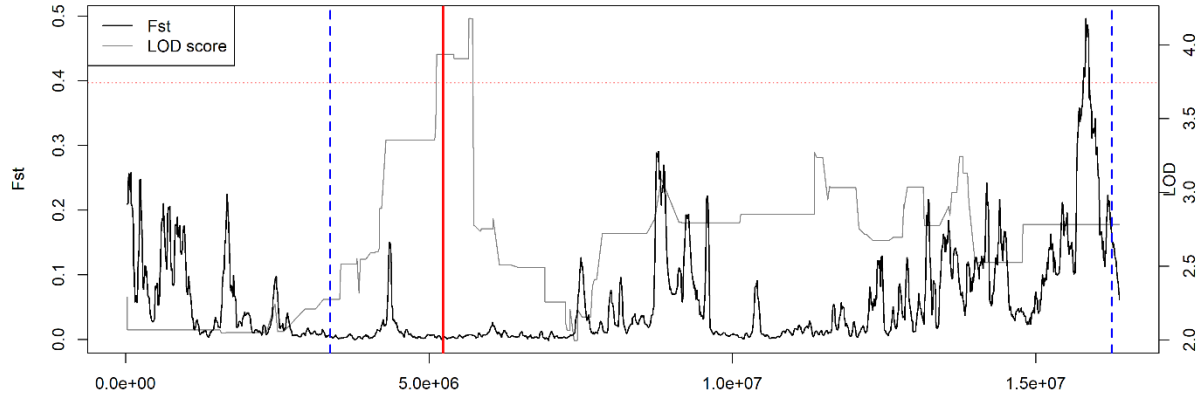
(Z)-11-Icosenylacetate

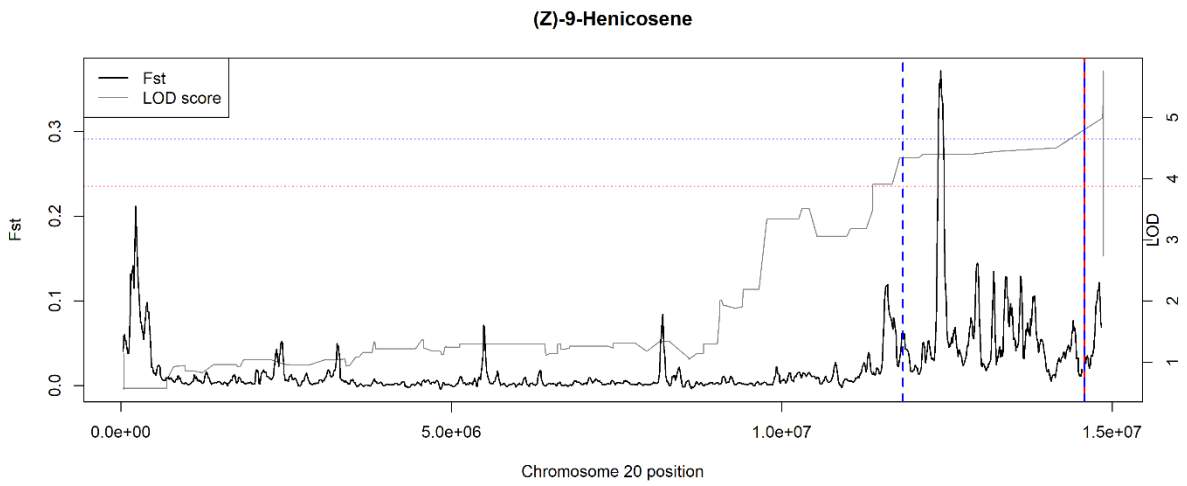


(Z)-11-Icosenylpropionate



Hexahydrofarnesylacetone




(Z)-9-Henicosene
Fst
LOD score
Fst
LOD
0.0
0.1
0.2
0.3
1
2
3
4
5
0.0e+00
5.0e+06
1.0e+07
1.5e+07
Chromosome 20 position

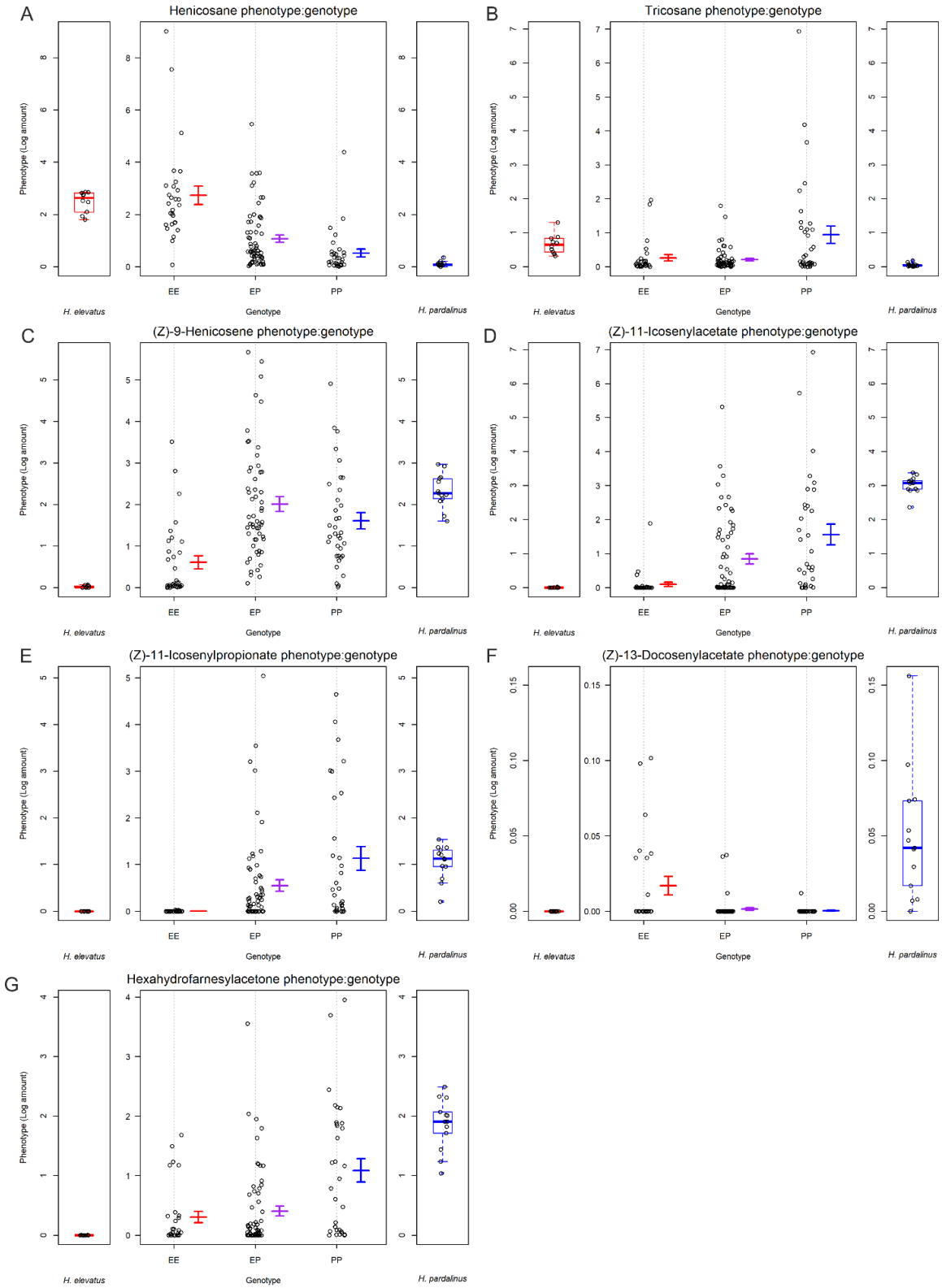
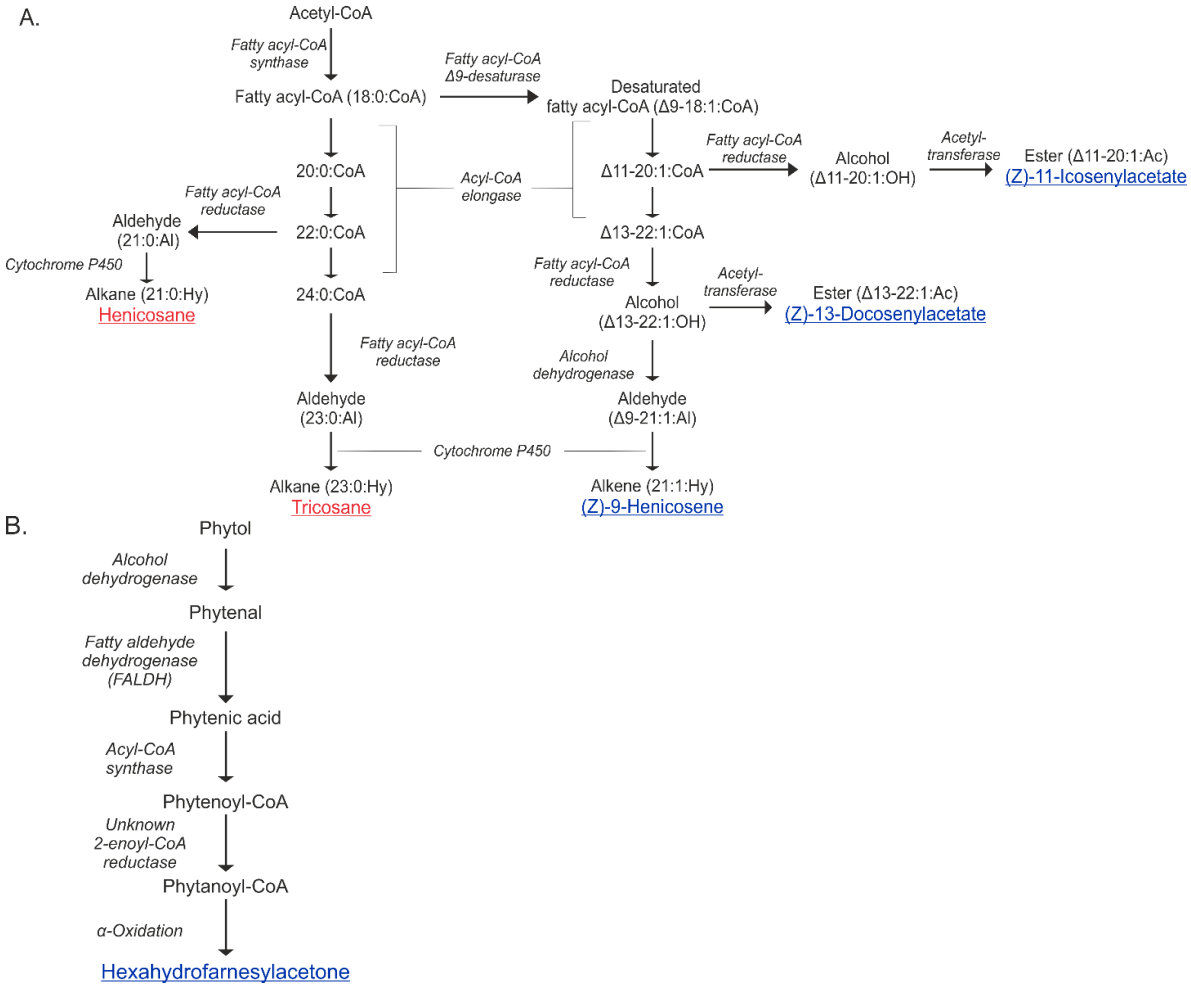


Figure 2.3 *A-G.* Phenotype:genotype plots for all compounds with significant LOD scores, alongside pure species phenotypes. For the phenotype:genotype plots, the mean phenotype is shown with standard error bars.

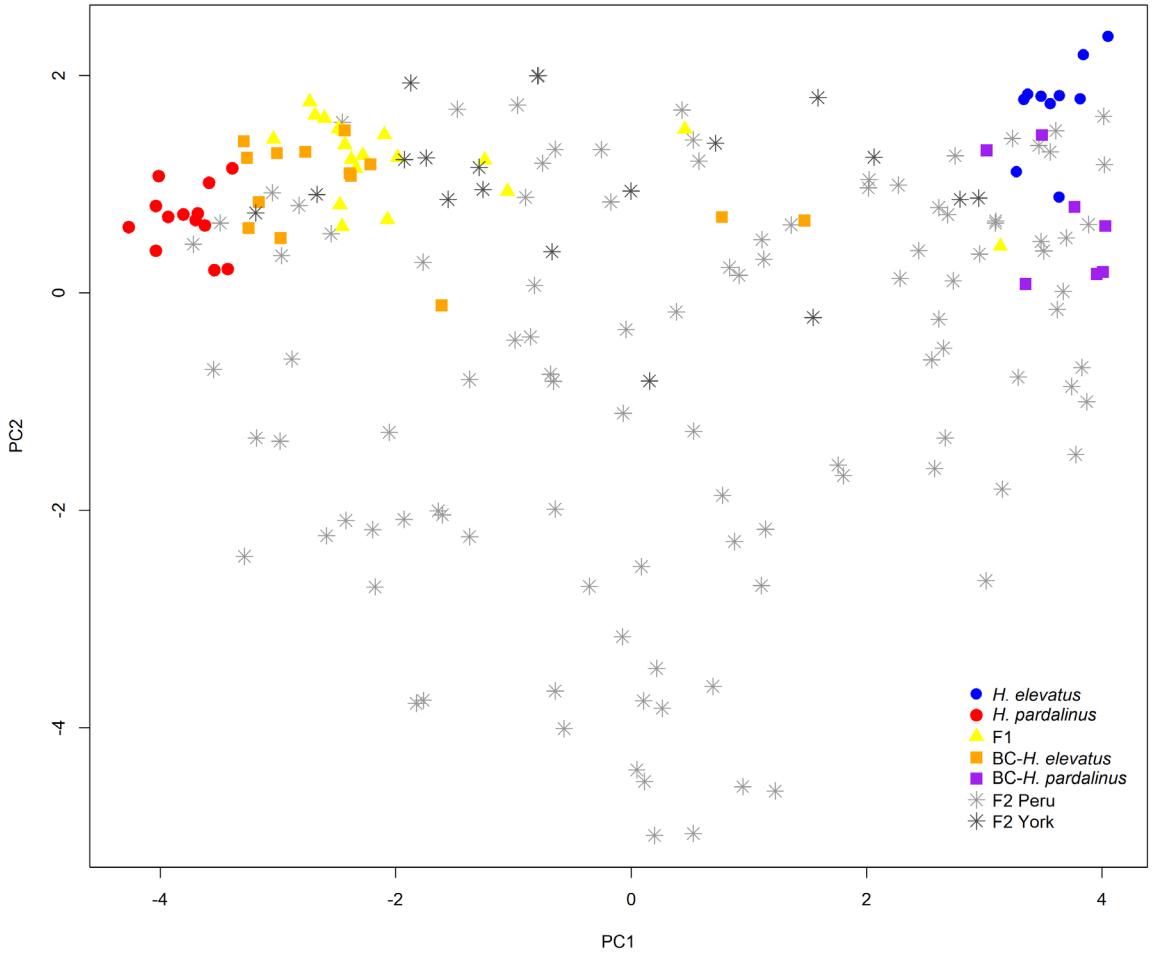
Compound	Class	Max LOD score	Chromosome	Physical peak	cM max	cM limits	Physical limits of the confidence interval	R2 (heritability)
Henicosane	Alkane	9.17	19	Hmel219001o:12805632	50.13	48.32-57.44	Hmel219001o:12591008 - 219002o:141794	0.29
Tricosane	Alkane	3.69	16	Hmel216002o:9583929	50.59	27.96-70.39	Hmel216002o:5395909 - Hmel216002o:10031338	0.13
(Z)-9-Henicosene	Alkene	5.76	20	Hmel220003o:14583363	88.13	83.78-93.5	Hmel220003o:11832016 - Hmel220003o:14584000	0.19
(Z)-11-Icosenylacetate	Ester	4.53	19	Hmel219001o:4738625	15.05	0-37.2	Hmel219001o:3680 - Hmel219001o:10514853	0.16
(Z)-11-Icosenylpropionate	Ester	3.98	19	Hmel219001o:2300117	4.55	0-16.88	Hmel219001o:3680 - Hmel219001o:4848438	0.14
(Z)-13-Docosenylacetate	Ester	4.88	16	Hmel216002o:10030821	66.42	60.82-81.32	Hmel216002o:10021318 - Hmel216002o:10031614	0.17
Hexahydrofarnesylacetone	Host plant	4.18	19	Hmel216001o:5261869	18.93	8.66-57.44	Hmel219001o:3373019 - Hmel219002o:141794	0.14

Table 2.1. Confidence interval details for each of the seven compounds that produced significant QTL peaks. Includes maximum LOD score of the QTL peak, the chromosome it is located on, its physical limits, its length in centiMorgans (cM), and theproportion of phenotype variance explained by the detected QTL.

Supplementary Material



Supp. Figure 2.1A. Putative biosynthetic pathway for the production of prominent *H. elevatus* (red) and *H. pardalinus* (blue) compounds. Shown here are only the compounds that produced significant LOD scores and could be mapped to specific QTLs, though this pathway can produce a much wider range of compounds (Mann *et al.*, 2017). Names of enzymes potentially involved in are shown in italics, based on (Reed *et al.*, 1994, 1995; Tillman *et al.*, 1999; Liénard *et al.*, 2014; He *et al.*, 2017; Mann *et al.*, 2017), and represent the types of proteins one may find within the QTL confidence intervals for these compounds. Fatty acyl-CoA reductases convert fatty acids into aldehydes with an alcohol intermediate which can then be converted into an ester, but the alcohol step is not always shown in the figure, unless esters are the relevant product. *B.* Pathway for phytol degradation into the *H. pardalinus* compound hexahydrofarnesylacetone based on (Van Den Brink *et al.*, 2005; Schulz, Yildizhan and van Loon, 2011), once again with relevant enzymes and reactions shown in italics.



Supp. Figure 2.2. PCA plot based on the 25 compounds that showed significant differences between *H. elevatus* and *H. pardalinus* (Rosser *et al.*, 2019), showing the two pure species, their respective backcrosses, the F1, and the F2 individuals bred in Tarapoto (Peru) and at the University of York (UK).

Compound	*H. elevatus*	Backcross to *H. elevatus*	F1	F2	Backcross to *H. pardalinus*	*H. pardalinus*
	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
Homovanillylalcohol	0.092 ± 0.128	0.874 ± 0.67	1.744 ± 1.413	7.573 ± 20.048	0.201 ± 0.246	0.223 ± 0.314
Hexahydrofarnesylacetone	**0 ± 0.001**	**0 ± 0**	**0.212 ± 0.192**	**2.14 ± 6.609**	**0.433 ± 0.617**	**5.997 ± 2.705**
Icosene	0 ± 0.001	0 ± 0	0.144 ± 0.128	0.01 ± 0.024	0.009 ± 0.016	0.041 ± 0.026
Henicosadiene	0.011 ± 0.02	0 ± 0	0.098 ± 0.05	0.013 ± 0.032	0.01 ± 0.015	0.034 ± 0.027
(Z)-9-Henicosene	**0.022 ± 0.027**	**0.054 ± 0.068**	**44.51 ± 14.911**	**13.562 ± 38.872**	**24.087 ± 17.576**	**10.087 ± 4.447**
Henicosane	**11.834 ± 4.428**	**5.728 ± 4.473**	**5.901 ± 3.503**	**89.237 ± 750.211**	**0.791 ± 1.099**	**0.115 ± 0.107**
Docosene	0 ± 0	0 ± 0	1.094 ± 1.262	0.117 ± 0.333	0.094 ± 0.075	0.122 ± 0.072
Oleyl acetate	0 ± 0	0 ± 0	0 ± 0	0.048 ± 0.373	0 ± 0	0.085 ± 0.044
Octadecyl acetate	0 ± 0	0.001 ± 0.003	0 ± 0	0.252 ± 1.841	0 ± 0	0.048 ± 0.023
Phytol	0.001 ± 0.004	0 ± 0	0.087 ± 0.089	0.071 ± 0.239	0.063 ± 0.04	0.109 ± 0.113
(Z)-9-Tricosene	0.045 ± 0.04	0.029 ± 0.023	8.171 ± 5.528	2.554 ± 7.19	1.246 ± 1.446	0.936 ± 0.744
Tricosane	**1.03 ± 0.672**	**0.394 ± 0.431**	**1.11 ± 2.01**	**9.637 ± 91.785**	**0.107 ± 0.093**	**0.053 ± 0.057**
(Z)-11-Icosenylacetate	**0.002 ± 0.007**	**0 ± 0**	**15.955 ± 13.725**	**15.437 ± 95.615**	**4.118 ± 3.319**	**20.089 ± 4.881**
Tetracosane	0.116 ± 0.072	0.074 ± 0.056	0.23 ± 0.347	1.132 ± 7.196	0.078 ± 0.057	0.034 ± 0.035
(Z)-11-Icosenylpropionate	**0 ± 0**	**0 ± 0**	**4.75 ± 3.97**	**4.505 ± 17.965**	**1.149 ± 1.438**	**2.032 ± 0.961**
Pentacosane	0.478 ± 0.384	0.266 ± 0.17	0.949 ± 2.224	10.726 ± 105.093	0.165 ± 0.073	0.208 ± 0.434
11-Methylpentacosane	0.539 ± 0.351	0.194 ± 0.142	0.356 ± 0.224	3.364 ± 20.922	0.181 ± 0.092	0.149 ± 0.114
(Z)-13-Docosenylacetate	**0 ± 0**	**0 ± 0**	**0 ± 0**	**0.005 ± 0.017**	**0 ± 0**	**0.052 ± 0.047**
Hexacosane	0.116 ± 0.051	0.082 ± 0.048	0.149 ± 0.433	1.132 ± 8.05	0.443 ± 1.56	0.053 ± 0.037
11-Methylhexacosane	0.033 ± 0.033	0.01 ± 0.009	0.014 ± 0.029	0.079 ± 0.737	0.202 ± 0.715	0.004 ± 0.007
Heptacosane	0.796 ± 0.363	0.716 ± 0.315	0.923 ± 0.672	15.156 ± 105.211	0.27 ± 0.139	0.17 ± 0.065
11-Methylheptacosane	0.124 ± 0.088	0.045 ± 0.031	0.061 ± 0.094	1.024 ± 6.942	1.424 ± 5.06	0.052 ± 0.018
Nonacosane	0.547 ± 0.183	0.703 ± 0.218	1.08 ± 0.838	9.853 ± 49.216	0.214 ± 0.109	0.235 ± 0.104
Octacosanal	0.449 ± 0.516	0.118 ± 0.111	0.277 ± 0.298	0.816 ± 2.886	0.248 ± 0.361	0.078 ± 0.123

Supp. Table 2.1. Comparison of mean amounts (nmol) of the 25 compounds that show significant differences between *H. elevatus* and *H. pardalinus*, across pure species, backcrosses, F1 and F2 hybrids. Compounds that produced significant LOD score in the QTL analysis, and could thus be mapped, are shown in bold.

Chapter 3: Sexual conflict and genital pheromone evolution in the Heliconiini tribe

Abstract

Antiaphrodisiacs are pheromones that are physically delivered from male to female individuals during copulation and discourage further mating events by rendering the female less attractive. These substances are often species-specific and their purpose is to reduce sperm competition between males. However, they do not always offer an honest signal of female receptivity. In polyandrous species, where females may be mated by multiple males, there is conflict of interests between the sexes which may lead to faster evolution of antiaphrodisiacs as the selective pressure to keep the signal stable is weaker. The Heliconiini clade of neotropical butterflies, which includes both a polyandrous and a monoandrous lineage, has been shown to produce a great variety of antiaphrodisiac blends with very low phylogenetic signal, hinting at rapid evolution. While it has been proposed that this may be due to male-male competition in polyandry, more recent studies have shown that *Heliconius* mating strategies are not strict and most species actually favour single matings. The aim of this study is to analyze and characterize the antiaphrodisiac blends of species from the Heliconiini tribe looking for shifts in evolutionary rates and any emerging patterns in antiaphrodisiac composition corresponding to shifts in mating strategy. I analyzed 36 species in total, including 24 previously uncharacterized species. I found no evidence for shifts in evolutionary rate corresponding to either shift in mating strategy, implying that male-male competition may have a weaker effect on pheromone diversification than previously thought. I also identified three prominent volatile compounds: (E)-β-ocimene, sulcatone and 4-hydroxycyclopent-2-en-1-one). Based on the function of (E)-β-ocimene as the behaviourally active antiaphrodisiac in *H. melpomene,* I propose a similar function in other species for the other two volatiles, and use *H. sara* to test this theory on 4-hydroxycyclopent-2-en-1-one.

Introduction

Antiaphrodisiacs in sexual conflict

Reproduction is not always a cooperative effort for all involved individuals. Sexual conflict is theoretically possible in any organism where multiple sexes are present, including documented examples in plants and fungi (Nieuwenhuis and Aanen, 2012; Lankinen and Green, 2015), and it is common in the animal kingdom due to clashing interests between mates as both males and females strive to control the outcome of the mating event (Arnqvist and Rowe, 1995; Chapman *et al.*, 2003; Hosken, Archer and Mank, 2019). This makes both sexes active parties in shaping the species' evolutionary processes, with often contrasting results as the sexes strives to evolve towards non-overlapping fitness optima (Arnqvist and Rowe, 2005; Parker, 2006; Hosken, Archer and Mank, 2019).

One way in which sexual conflict takes place is via male control over the frequency of female mating events. In many species it is advantageous for females to mate multiple times and with different males (polyandry) (Forsberg and Wiklund, 1989; Arnqvist and Nilsson, 2000): repeated mating events offer females the chance to copulate with more fit males they may encounter after the initial mate, and competition between multiple partners may lead to increased offspring fitness as the genetically preferrable sperm is more likely to survive and fertilize the ova (Jennions and Petrie, 2000). Moreover, in many species polyandry also leads to greater fertility and increased offspring production at the fully offset cost of slightly reduced longevity (Arnqvist and Nilsson, 2000). However, this leads to male-male conflict which is deleterious for individual males due to the phenomenon of sperm competition. The latter takes place within a multiply-mated female between sperm cells of several males in order to fertilize the ova (Simmons, 2001; Parker, 2006). This process is widespread in living kingdoms and among animal phyla, including vertebrates (Birkhead, 1998; Gomendio et al., 1998) and molluscs (Squires *et al.*, 2015). However it is best known in insects, where it can be resolved with a variety of outcomes depending on the species (reviewed in Parker, 2006); while in some cases the result is mixed progeny, in cases such as twice-mated *Drosophila* the last male benefits from sperm precedence over previous males (Price et al., 1999; Smith et al., 2017), albeit this effect is controlled by multiple factors and diluted when the female is able to mate more than twice (Laturney et al., 2018). Last male sperm precedence is also seen in the Orthoptera (Parker, 2006) as well as in *Tenebrio molitor* (Siva-Jothy *et al.*, 1996), several Lepidopteran species (Solensky & Oberhauser, 2009; Xu and Wang, 2010; Yan *et al.*, 2013), birds (Birkhead, 1998) and cephalopods (Gomendio et al., 1998), among others. Meanwhile, other species, including parasitoid wasps (Ablard *et al.*, 2014), social spiders (Jones and Parker, 2008) and newts (Jones et al., 2002), have developed adaptations to ensure first-male sperm precedence, often involving extreme solutions: for example, in the mosquito *Aedes aegypti* the first male to copulate with the female is also the one whose sperm will fertilize most ova, as monogamy is enforced

via a male-produced protein-based substance ("matrone") that renders the female sterile for life (Craig, 1967; Fuchs et al., 1968; Fuchs et al., 1969).

Other species have evolved different behaviours such as mate guarding and sperm plugs in order to overcome intrasexual conflict and nullify sperm competition (Simmons, 2001; Parker, 2006). Antiaphrodisiac pheromones are one such strategy. Originally discovered in the beetle *Tenebrio molitor* (Happ, 1969), they were later described in multiple other taxa, including one vertebrate, the garter snake (Ross and Crews, 1977, 1978), but mainly other insects such as *Drosophila* (Tompkins and Hall, 1981; Scott, 1986; Laturney and Billeter, 2016; Smith et al., 2017), true bugs (Brent and Byers, 2011), sweat bees (Kukuk, 1985) as well as many other species of Hymenoptera (Ayasse et al., 2001), and butterflies (Andersson *et al.*, 2000; Andersson *et al.*, 2003; Schulz *et al.*, 2008).

Like all pheromones, antiaphrodisiacs are mixtures of compounds (or occasionally single compounds) that elicit a behavioural and/or physiological response in a receiving individual (Karlson and Lüscher, 1959). The specific function of antiaphrodisiacs is to inhibit sexual activity by rendering females less attractive to other males (Happ, 1969). The purpose of this, as was hypothesized in *Tenebrio,* is for the male to ensure that the female does not mate again, ensuring his paternity over that of a competing male (Happ, 1969).

Antiaphrodisiacs may or may not play a part in sexual conflict, depending on the mating strategy employed by the species in question; antiaphrodisiac-producing males always have an investment in avoiding sperm competition, but female interest in repeated mating events varies between species as well as within species (based on factors such as time elapsed since the latest copulation) (Andersson, *et al.*, 2000; Andersson *et al.*, 2003; Estrada *et al.*, 2011). In situations of monoandry, where females carry offspring from a single male at a time, these pheromones provide a honest signal of female receptivity (Ross and Crews, 1978; Andersson *et al.*, 2000; Estrada *et al.*, 2011). This is because mating events come at a cost. While both sexes have to endure the inherent time and energy expenditure of reproduction (Daly, 1978), female reproductive costs come with the greater vulnerability to predation while engaged in courtship or copulation as well as the greater risk of being exposed to parasites or diseases from interaction with multiple individuals, and the potential of being injured or killed by the male (Daly, 1978; Watson et al., 1998). Thus, when frequent mating has a negative impact on female reproductive success, or when females are otherwise not physiologically receptive to multiple copulation events, the two sexes have non-opposing goals and sexual conflict does not occur. It has been hypothesized that this may lead to evolutionary stabilization of the antiaphrodisiac signal as it benefits both males and females and is under no selective pressure to change (Estrada et al., 2011).

Sexual conflict arises in situations where females become ready to mate before the antiaphrodisiac signal has dissipated, be it because the species itself is polyandrous or because enough time has elapsed since the first mating. In this case, antiaphrodisiacs become a dishonest signal, since they reflect the male's interests but not the female's (Estrada et al., 2011). For example, in Drosophila, male antiaphrodisiacs are detrimental to females (Wigby and Chapman, 2005; Laturney and Billeter, 2016; Smith et al., 2017). Female Drosophila, which strongly favour sperm from later copulation events, actually go to the metabolic trouble of producing an anti-antiaphrodisiac pheromone that renders them attractive again and counteracts the action of the first male's antiaphrodisiac (Laturney and Billeter, 2016). Under sexual conflict, it is predicted that the signal may be under less pressure to remain stable, potentially leading to rapid evolution and disruptive selection (Arak and Enquist, 1995).

Heliconius antiaphrodisiacs

Heliconius butterflies are well known models in evolutionary biology, and they provide a convenient system for the study of pheromone evolution. Mated *Heliconius* females carry a bouquet of male-contributed odours (Gilbert, 1976) whose antiaphrodisiac function was first demonstrated in *H. melpomene* by Schulz *et al.* (2008). *Heliconius* antiaphrodisiacs are produced by scent glands at the tip of the abdomen, and in *H. melpomene* they are comprised of a prominent volatile compound, (E)-β-ocimene, accompanied by a complex matrix of large, less volatile fatty acid esters (Schulz *et al.*, 2008; Vanjari *et al.*, 2015). This matrix has little behavioural effect on its own, at least in *H. melpomene,* but it seems necessary to modulate the rate of (E)-β-ocimene evaporation, enhancing its repellent effects (Schulz *et al.*, 2008). Meanwhile, (E)-β-ocimene is the behaviourally active compound, as it renders unmated females less attractive to males when topically applied to the tip of their abdomen (Schulz *et al.*, 2008).

Heliconius is considered a suitable model for researching antiaphrodisiac evolutionary dynamics because of an early split, 11.8 Ma (Kozak *et al.*, 2015), into two lineages with two different mating strategies and different rates of male-male competition, allowing for phylogenetic contrasts between a "monoandrous" and "polyandrous" system. The two lineages are traditionally referred to as pupal-mating (comprising two clades, *erato* and *sara*) and free-mating (comprising the *melpomene* and silvaniform clades alongside the smaller *aoede, doris* and *burneyi* clades) (Fig. 3.1). (Kozak *et al.*, 2015). All pupal-mating species descend from the same common ancestor (Beltràn *et al.*, 2007), and in this lineage, males guard female pupae and mate as soon as they commence eclosion, or as soon as they have fully emerged, thus male-male competition is low and repeated mating events are measurably infrequent in the wild (though not completely unheard of) (Gilbert, 1976; Deinert *et al.,* 1994; Estrada *et al.*, 2010). In the free-mating lineage, where butterflies mate as fully emerged free-flying adults,

repeated mating is more common than in pupal mating species (Walters *et al.*, 2012) and male-male competition is more prominent.

Estrada et al. (2011) were the first to investigate antiaphrodisiac macroevolution in *Heliconius* butterflies. They characterized the antiaphrodisiac blends of 11 species within this genus, in order to test whether these pheromones evolved at an increased rate in species belonging to the polyandrous lineage compared to the monoandrous one, with the prediction that increased sperm competition in the polyandrous lineage would lead to more sexual conflict and thus more rapid antiaphrodisiac diversification due to the increased pressure to keep females from remating (Estrada *et al.*, 2011), as first described in (Arak and Enquist, 1995). They found that antiaphrodisiacs carry low phylogenetic signal across the genus, hinting at low similarity in blend composition even between closely related species, and that species within the polyandrous lineage show more divergent blends than species within the monoandrous lineage where male-male competition is less frequent, as they predicted (Estrada *et al.*, 2011).

Since Estrada et al. (2011), some studies have challenged the notion of the two *Heliconius* lineages as strictly monoandrous and polyandrous, a caveat originally suggested by Estrada et al. themselves. Walters et al. (2012) demonstrated that while free-mating females are much more likely than pupal-mating females to carry sperm from more than one male, the most common scenario is still single-mating regardless of the lineage. Mating events in the free-mating clade also tends to be separated by long periods of time, which in itself reduces sperm competition. It has also been shown that pupal mating is a facultative behaviour, as species from pupal-mating species do occasionally mate later in life (Walters *et al.*, 2012), and it is actually less frequent than free mating in the *erato* clade, and more frequent than pupal mating, although it it's currently unknown whether this also applies to species from the *sara* clade (Thurman *et al.*, 2018).

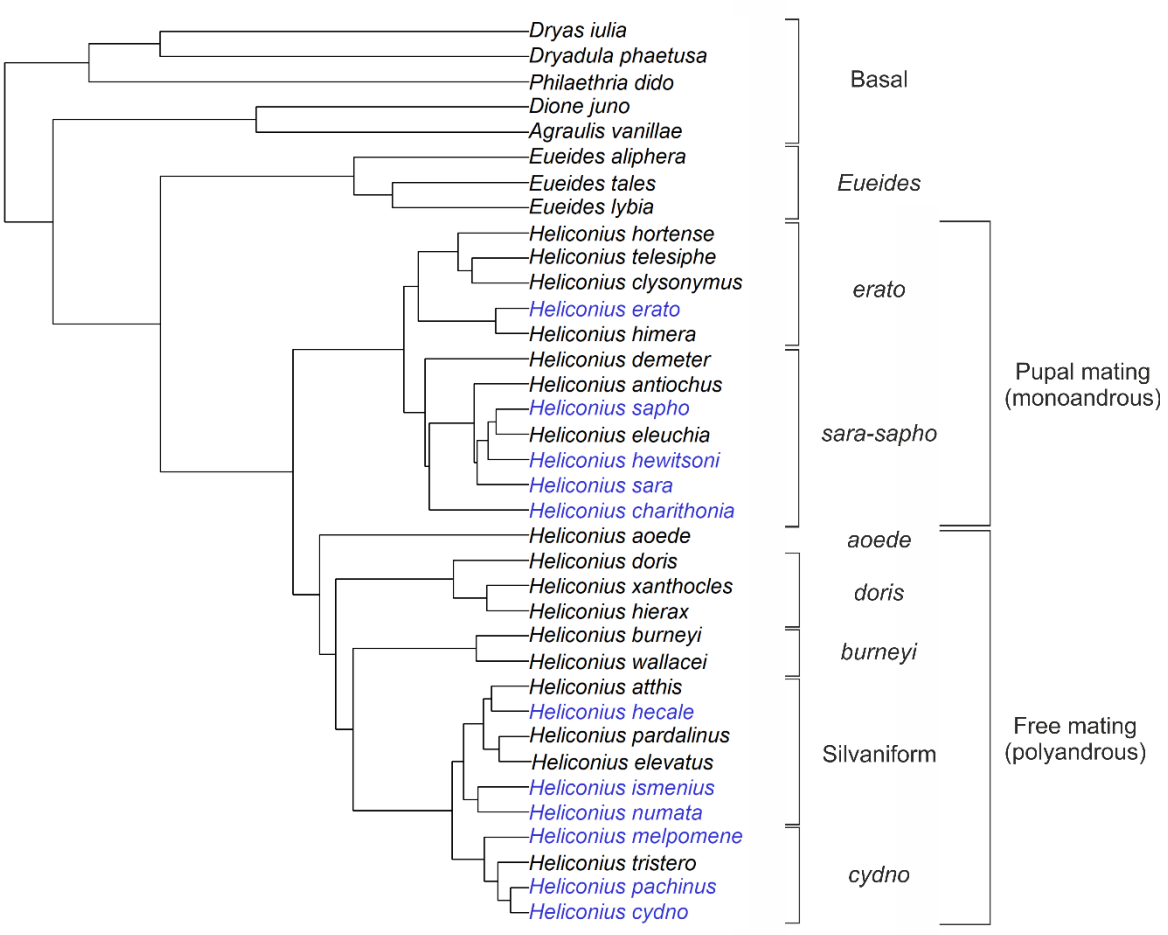
With this additional knowledge, with the aid of recently developed techniques for multivariate phylogenetic analyses, I test here the hypothesis presented by Estrada et al. (2011) on a greater variety of species than they studied, from both lineages, seeking to determine whether the relatively weak extent of male-male competition in *Heliconius* is enough to promote greater antiaphrodisiac diversification in the free- mating clade. This includes fitting models accommodating for shifts in evolutionary rates at different points in the phylogeny, methods to compare evolutionary diversification rates of different subgroups and testing within-species similarity of the pupal- mating lineage against the free- mating lineage. Additionally, I profile the antiaphrodisiac blend of 24 previously uncharacterized species paying particular attention to volatile compounds which may be more likely to have an antiaphrodisiac effect.

Materials and Methods

Specimen collection

Butterflies used in this study were mostly obtained from the wild in Panama and in Ecuador, the latter collections being made by Stephanie Ehlers (Braunschweig TU) and Chris Jiggins (University of Cambridge), or from captive populations from the Smithsonian Tropical Research Institute (STRI) insectaries in Gamboa (Panama). The main collection sites for wild butterflies were Soberania National Park and the Reserva Forestal El Montuoso (Panama). *H. eleuchia* and *H. timareta* were collected and dissected by K. Darragh (University of Cambridge), and data collected from these samples hashave already been published (Darragh *et al.*, 2020). *H. antiochus, H. atthis, H. charithonia, H. ismenius, H. himera* and *H. hortense* were sampled from pupae obtained from Stratford Butterfly Farm (Stratford, United Kingdom), from populations raised in Ecuador, Colombia and El Salvador (Fig. 3.1).

Figure 3.1. Heliconiini phylogeny adapted from Kozak et al. (2015), showing all sampled species for this study. Species previously featured in Estrada (2011) are shown in blue, whereas new additions are shown in black. For a full depiction of the phylogeny with the sampled species, and the number of individuals sampled for each species, see Supp. Fig. 1.1.



Tissue extraction and preparation

Wild butterflies were sacrificed at most 2 days following capture. Butterflies obtained as pupae were sacrificed as soon as they reached sexual maturity, 5-10 days post-emergence. The tissue samples were extracted immediately following the butterfly's death. The tissue extracts consisted of the last abdominal segment in both males and (when available) female samples. In males, this is where structures known as claspers are located, which carry the relevant pheromone producing glands (Eltringham, 1925; Schulz *et al.*, 2008; Estrada *et al.*, 2011). Females on the other hand possess complementary structures to the claspers where these scent bouquets are deposited (Eltringham, 1925, Estrada *et al.*, 2011). The samples were extracted using clean sterile forceps and scissors and immediately suspended in 200µL of dichloromethane (containing 1 ng/µL 2-acetoxytetradecane to act as an internal standard) in a 2 mL glass vial.

Gas Chromatography-Mass Spectrometry settings

Abdominal tip extracts from both sexes were analyzed on a GC-MS system which consisted of a 7890A GC-System coupled with an MSD 5975C Mass Analyzer (Agilent Technologies, Santa Clara, USA) fitted with a HP-5MS column (50m length 0.25 mm I.D., 0.25 µm film thickness; Agilent Technologies). The ionization method was electron ionization with an electron energy of 70 eV. The instrument conditions were the following: inlet pressure 9.79 psi, He 20 mL min^{-1}, injection volume 1 µL. The GC was programmed with the following temperature ramp: start at 50°C, increased by 5°C min^{-1} to a maximum temperature of 320°C, then held for 5 min for a total run time of 69 min. Most samples were analyzed in the Chemical Ecology laboratories at the STRI Earl S. Tupper Center in Panama City (Panama) in 2018. *H. timareta* and *H. eleuchia* samples were run by Darragh K., with the same method and on the same GC-MS system, between late 2017 and early 2018. A subset of samples was run with using the same settings conditions and on the same instrument model at Braunschweig Technische Universität in Braunschweig (Germany), and some of these samples were analyzed using a shorter temperature ramp, starting at 50°C, then increasing at 10°C min^{-1} to a maximum temperature of 320°C for a total run time of 45 min.

Samples from the *Heliconius sara* experiment were run at the University of York Department of Chemistry (York, UK) on a GC-MS system which consisted of a 7890A GC-System coupled with a Waters GCT Premier TOF Mass Analyzer (Waters Corporation, Milford, MA, USA) fitted with a Phenomenex ZB5-MSplus (30m x 0.25mm x 0.25µm) column (Phenomenex, Macclesfield, UK). For improved mass accuracy, MS grade perfluorotributylamine (PFTBA a.k.a. Heptacosa, Code:PC0568; Apollo Scientific Ltd., Stockport, UK) was utilized as a constant-flow calibrant within the mass analyzer. The ionization-method was electron impact with a collision-energy of 70 eV. The conditions were the following: inlet

pressure 9.79 psi, He 20 mL min^{-1}, injection volume 1 µL. The GC was programmed with the same 69 min temperature ramp used for the Panamanian samples.

All batches of samples were run alongside a C6-C40 alkane ladder prepared at Braunschweig TU (Braunschweig, Germany). Said alkane ladder was used to calculate the retention index (RI) for each compound, necessary for identification and comparison with existing databases, particularly with the NIST Mass Spectra and Retention Indices databases (Linstrom and Mallard, 2021).

Data processing

The composition of Heliconiini abdominal extracts is very rich and complex, thus I opted to process the GC-MS data via non-targeted analysis using the Python package PyMassSpec (Davis-Foster, 2019, originally from O'Callaghan *et al.*, 2012). GC-MS peak deconvolution and alignment of different sample chromatograms was carried out largely in PyMassSpec, however due to the differences between the STRI and the Braunschweig TU instruments, different subgroups of samples could not be aligned within PyMassSpec and required a different method to minimize batch effects. Alignment between batches was thus based on the two most intense *m/z* fragments signals in each peak, and on the retention index (RI) of detected peaks. For two peaks to be aligned and processed further as a single compound, the conditions were complete correspondence between the two largest *m/z* signals, as well as similar RIs (±0.8% of the RI value), based on the range of RIs observed in cholesterol (a ubiquitous and late-eluting compound) and thus needed to align the latter among different sample types. This procedure was carried out via a Perl script, and was used to align 3 data from samples analyzed in three different batches: Panama (PAN), including both the 2018 samples and the *H. eleuchia/timareta* samples from Darragh (2020), Braunschweig 69 min gradient program (BWL), Braunschweig 45 min gradient program (BWS). The resulting dataset was manually checked for alignment errors via the NIST AMDIS (Automated Mass Spectral Deconvolution and Identification System) software. Species collected by Ehlers, S. (Braunschweig TU) in Ecuador were placed in their own dataset, designated as ECU. An additional batch of samples, run in 2012 at the University of Braunschweig, consisting of samples from *H. elevatus* and *H. pardalinus butleri*, was merged manually and designated as BWF.

Compounds were scored between a minimum RI=900 and a maximum RI=3077, corresponding to cholesterol in this dataset, consistently with a previous report of cholesterol with the same column type (Steiner et al., 2005, reported RI=3075). Relative proportions of each compound were calculated from the peak areas and expressed as percentages of the sum of all peak areas. Since a very large number of peaks were detected (several thousands across all datasets), most of which rarely contributed more than 0.5% of the total blend of any individual, a hard threshold was used to reduce

the complexity of the dataset. For each sample, any peaks whose relative amount fell below the 75% percentile were set at zero, and every peak whose relative amount never exceeded the 75% percentile was excluded from the analysis. For the five batches of samples, this threshold resulted in the removal of any peaks that contributed less than 0.18%, 0.13%, 0.27%, 0.11% and 0.47% of an individual's total blend in the PAN, BWL, BWS, BWF and ECU batches respectively. Rare compounds that appeared less than 5 times across the entire dataset of 157 samples were also excluded from the analysis. Unknown compounds were identified by in the Stefan Schulz's laboratory at Braunschweig TU.

Chemical dissimilarity between samples was obtained with the 'vegdist' function in R(vegan) (Oksanen *et al.*, 2019), and expressed as Euclidean distance. The data was represented as a 5-axes NMDS plot computed by the 'metaMDS' function (also from R(vegan)). Euclidean distance was also used to calculate species and clade contributions to pheromone diversity via a permutational analysis of variance (PERMANOVA/ADONIS) based on 10,000 permutations. The purpose of this was to assess blend similarity within species and between species. Average Euclidean distance between conspecifics and within-species dispersion (distance from the species' median pheromone composition) (Anderson, 2006; Anderson *et al.*, 2006) were used to compare interspecific differences between the pupal- and free-mating lineages in a Welch two-sample t-test, to demonstrate whether the free-mating species, due to increased sexual conflict, would show more differences between conspecifics. Both parameters were obtained via R(vegan).

In androconial extracts (see Chapter 1), dissimilarity in chemical profiles is greatly enhanced in sympatric sister species pairs, an important aspect of their potential role as reproductive barriers in the Heliconiini. Antiaphrodisiacs, not being involved in courtship rituals, are poor candidates as pre-zygotic reproductive barriers. To validate this assessment, I calculated the average dissimilarity between the 10 sister species pairs in the dataset: *Dryas-Dryadula, Eueides lybia-Eueides tales, H. clysonimus-H. telesiphe, H. erato-H. himera, H. hierax-H. xanthocles, H. ismenius-H. numata, H. atthis-H. hecale, H. elevatus-H. pardalinus* and *H. cydno-H. pachinus.* I obtained range overlap information from Rosser *et al.* (2015) and used a glm model to test for the effect of overlap on genital extract dissimilarity. I used branch length (divergence time in Myr as extracted from (Kozak *et al.*, 2015)) to control for phylogenetic effects as less recently diverged species are expected to show a higher degree of divergence regardless of range overlap.

While the analysis was non- targeted, any necessary compound identification was carried out in the AMDIS (Automated Mass Spectral Deconvolution and Identification System) software (Mallard and Reed, 1997; Stein, 1999), distributed by NIST (National Institute of Standards and Technology). Libraries compiled at Braunschweig TU (Braunschweig, Germany) were used for compound

identification alongside data compiled on the NIST databases. All analyses were run in R using the base stats package. Where needed, the dataset was split into early-eluting (RI≤1800) and mid/late-eluting compounds (RI>1800). The value of RI=1800 was chosen based on a gap in retention indices between the two categories of compounds (Supp. Fig. 3.1): in fact, the density of retention indices detected in the dataset increases rapidly past this point, reflecting the high amount of different compounds detected at each retention index window in the mid/late-eluting region (Supp. Fig. 3.1). The latters make up the "matrix".

Phylogenetic analysis

All phylogenetic analyses made use of the published Heliconiini phylogeny from Kozak et al. (2015), based on 22 mitochondrial and nuclear markers. Analyses were carried out at the species level, and one race was chosen for species where more than one race was available in order to avoid polytomies; *H. m. plesseni* for *H. melpomene* and *H. e notabilis* for *H. erato.* These races were chosen since their branch lengths were known from the Kozak et al. (2015) phylogeny.

To test for variation in rate of genital blend diversification, rate shift evolutionary models were fitted using two functions from two different R packages: transformPhylo.ML from R(motmot) with a "tm1" model (Puttick *et al.*, 2020) and mvSHIFT from R(mvMORPH) (Clavel *et al.*, 2015). The motmot is used for fitting models with no *a priori* expectations for the location of rate shifts. The minCladeSize parameter, denoting the minimum number of species for a subclade to be considered as a candidate for a rate shift, was set to 3 to counter the possibility of rate shifts appearing near the tips, for example between sister species pairs. As the tm1 model returns an AICc score for each additional shift it locates as well as for the base no-shift Brownian model, it was run alongside the calcCutOff function. The latter returns the minimum AICc cutoff to evaluate the best fitting model based on 1000 permutations.

Meanwhile, mvSHIFT from mvMORPH fits and allows for easy comparison of models where the point of the shift is set *a priori,* either at specific nodes in the phylogeny or at set time points. The mvSHIFT function can also be used in combination with the simpler mvBM and mvEB functions, that set simpler models of Brownian motion (no burst) and Early-Burst respectively. The tested models were as follows: Brownian motion to Early-Burst (BM-EB) with shift at the free-mating lineage, Brownian motion with rate shift (BMM) at the free-mating clade, BMM at the pupal-mating clade, EB-BM at the pupal-mating clade (also known as acceleration-deceleration or ACDC model), BM-EB with a shift at the advanced lineage (corresponding to the common ancestor of *Eueides* and *Heliconius*), simple BM and simple EB. All these models were tested separately on the full blend, on early eluting compounds (RI≤1800) and on mid/late-eluting compounds. The same models were also tested on a sub-tree including only *Heliconius* to filter out the effect of a potential evolutionary burst at the

Eueides/Heliconius node, which would make detection of free-/pupal-mating rate shifts more difficult. These models were fitted? to the 5 NMDS axes rather than the relative amounts of compounds.

Evolutionary rates of the pupal- and free-mating clades were compared using the compare.evol.rates function from R(geomorph) on logged relative amounts of the compounds, with significance estimated from 1000 random permutations (Adams and Otárola-Castillo, 2013). The rates were also compared among smaller *Heliconius* sub-clades as per Kozak (2015), as follows: *erato, sara-sapho, doris, cydno* and silvaniform, in case any shift in rate could be detected at that level rather than at the pupal/free-mating split. The same test was also repeated on a smaller dataset including only the species analyzed by Estrada et al. (2011) in order to compare the results. Multivariate phylogenetic signal was calculated on relative amounts of compounds for the whole blend as well as the two subsets using the K_{mult} statistic described in (Adams, 2014) with 10000 random permutations.

Phylogenetically corrected correlations between compounds were calculated using R(phytools) (Revell, 2012) as detailed in (Revell, 2017). The resulting p-values and Pearson correlation coefficients were used to plot a correlation matrix via the R(corrplot) package (Wei and Simko, 2017). This was done to test for any biosynthetic or genomic linkage between different compounds, implying a potential signalling benefit, such as behavioural synergies, in specific compound combinations.

Early eluting compound occurrence in *Heliconius sara*

The qualitative and phylogenetic analysis on wild-caught individuals from Panama indicated that *Heliconius sara*'s signature early eluting compound is 4-hydroxycyclopent-2-en-1-one, but its behavioural function has not been tested in *Heliconius*. It is sequestered from the host plant and derives from cyanogenic glucosides involved in *Heliconius* toxicity (Pinheiro de Castro *et al.*, 2019), so it may accumulate ubiquitously in the butterflies' tissues, a characteristic that would make it an unlikely candidate for an antiaphrodisiac function. If, however 4-hydroxycyclopent-2-en-1-one has an antiaphrodisiac function like (E)-β-ocimene, it would be expected to appear in males and mated females, but not in virgin females.

To further investigate the potential antiaphrodisiac role of 4-hydroxycyclopent-2-en-1-one, additional *H. sara* were obtained from experimental stocks maintained at the University of Sheffield by Nicola Nadeau. Genital samples were collected from 13 individuals including 4 virgin females, 3 mated females, 3 young males (0-1 days post eclosion) and 3 older, sexually mature males (8+ days old). Following results from the first batch of samples, wherein the compound of interest was never detected in stark contrast with wild Panamanian individuals, a host plant assay was carried out to test whether the presence or absence of 4-hydroxycyclopent-2-en-1-one was due to diet of the larvae. For this experiment, samples were collected from 12 sexually mature individuals in total, that had been

subjected to two different feeding regimes as larvae; 3 males and 3 females were fed on *Passiflora biflora* and the remaining individuals, 3 males and 3 females, were fed on *Passiflora auriculata.* Of these two host plant, *P. biflora* is the suboptimal host for *H. sara,* which prefers *P. auriculata* in the wild (Brown, 1981), so the former may be expected to produce an altered chemical profile while the latter may produce one closest to the wild *H. sara* signature. From male plant assay individuals, I collected androconia, genitals and a control region of the wing, whereas from females I only collected androconia and genitals (note that female *Heliconius* do not have androconial structures, but the equivalent region was extracted nonetheless), making up a total of 30 samples across all tissues, sexes and host plants. All of these samples were collected and prepared by O'Roarty, H. and Scott, E. at the University of Sheffield Nadeau lab. For each sample, I scored the amount of 4-hydroxycyclopent-2-en-1-one, alongside a structural compound found in all areas of the butterfly wing, nonacosane, to act as a control for fluctuations in the amount of sampled tissue. Androconia and control samples were collected due to past results which found 4-hydroxycyclopent-2-en-1-one to be prominent in Panamanian *H. sara* male androconia (Chapter 1), however in this chapter I will be presenting results strictly pertaining to the genitals. I used Welch's two-sample t-test to first test for differences in amounts of 4-hydroxycyclopent-2-en-1-one between *P. auriculata* genitals and *P. biflora* genitals (males and females together), then to test for differences between male and female genitals within the *P. auriculata* group.

Results

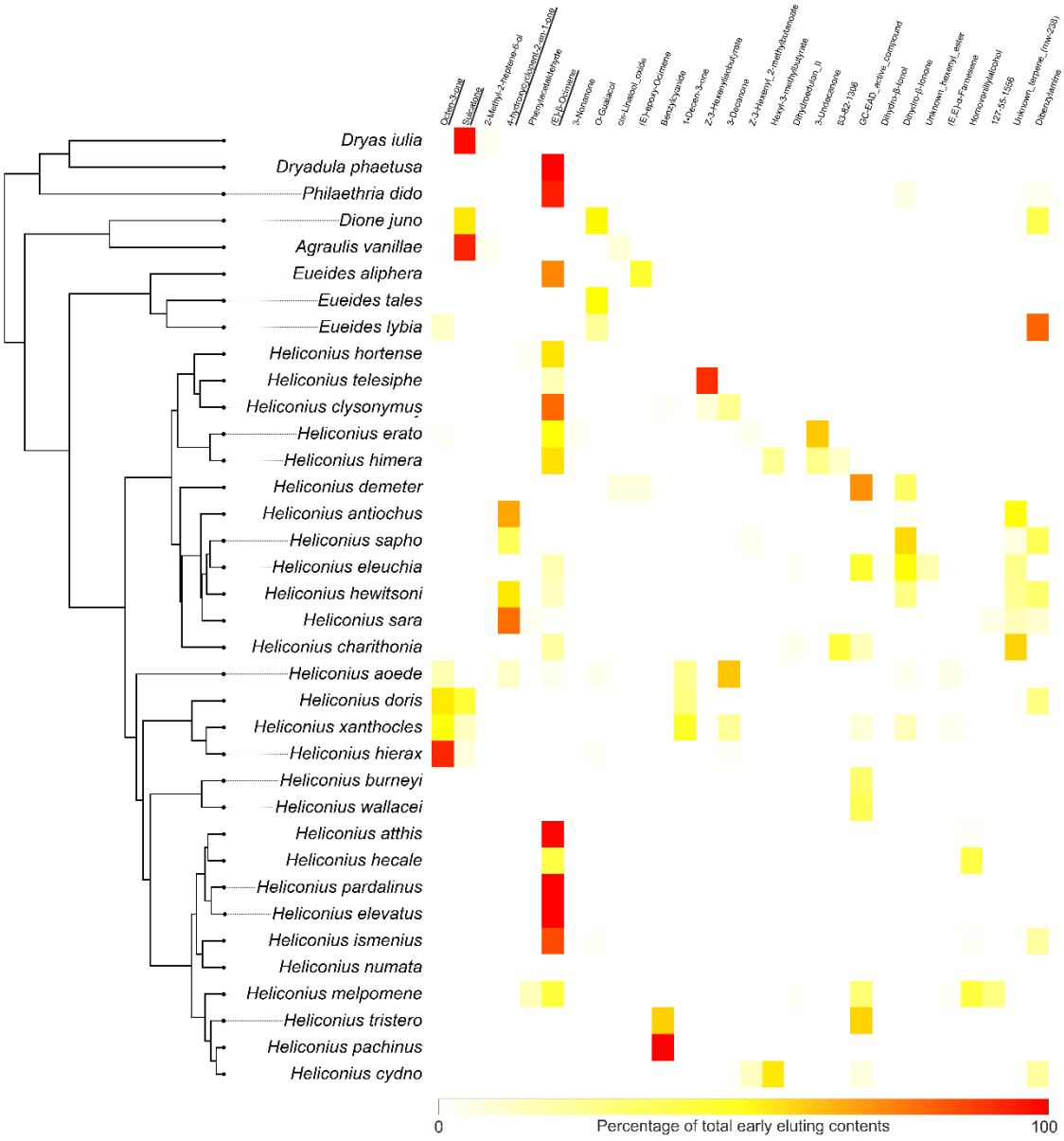
Qualitative analysis of genital contents

An exploratory analysis of the dataset was the first step in this study. Thousands of compounds were detected via PyMassSpec non-targeted analysis, requiring harsh filtering measures to exclude all compounds present in very low amounts. Thus, I opted to exclude all compounds whose amounts never exceeded the 75% percentile value for each dataset. After the percentage threshold was applied, I retained 441 compounds across the 157 samples from the 36 different species featured in the study across the five filtered datasets PAN (Panama), BWS (Braunschweig short temperature ramp), BWL (Braunschweig long temperature ramp), BWF (Braunschweig from (Mann *et al.*, 2017)) and ECU (Ecuador) were merged. A large portion of compounds appeared fewer than 5 times in the dataset, but 215 compounds appeared 5 or more times. The data was acquired from two GCMS instruments (albeit the same model), with two slightly different temperature ramps, and over long periods of time, which required a protocol for aligning outputs from different sources. As samples were run in batches alongside an alkane ladder, the latter could be used to calculate the retention index (RI) for each GC peak, a fixed parameter which was used alongside mass spectrometry

information to align peaks from different datasets. This method successfully reduced the batch effect seen among the five different datasets, albeit further adjustments had to be performed manually.

Having assembled the full combined dataset, I first observed general trends in the composition of genital contents. Much like in the androconia (Mann *et al.*, 2017; Chapter 1), most species showed a combination of fatty acid derivatives and potential host plant derivatives, albeit some of these seemingly plant-derived compounds may actually be produced endogenously by the butterflies, as is the case for a previously described *H. melpomene* antiaphrodisiac compound, (E)-β-ocimene (Darragh, Orteu, *et al.*, 2019). The majority of species show the same pattern as that previously reported in *H. melpomene* (Schulz *et al.*, 2008) with one or a few early-eluting compounds present in large amounts, usually followed by a gap, then by the matrix of fatty acid esters (Schulz *et al.*, 2008). This same pattern had also been reported in some of the species featured in (Estrada *et al.*, 2011). Of all 36 species present in this study, only 6 did not show this pattern: the only four where early-eluting compounds were not present in detectable amounts were *H. burneyi, H. wallacei, H. numata* and *H. cydno,* and only very low abundances of some early-eluting compounds were detected in *H. demeter* and *H. charithonia.* As early eluting compounds were a major point of interest, their identification took priority over that of mid/late-eluting compounds and I investigated their patterns of occurrence throughout the featured species. Among the early eluting compounds, three are particularly notable due to either being widespread or to being exclusive to specific clades: (E)-β-ocimene, found in the *erato, cydno* and silvaniform *Heliconius* clades as well as in *Eueides aliphera, Dryadula phaetusa* and *Philaethria dido*, sulcatone, found in all other basal species in the study and in *H. doris*, 4-hydroxypent-en-2-one, found in the *sara* clade exclusively, and octen-3-one, found in all species of the *doris* clade and in low amounts in *H. aoede* and *E. lybia* (Fig. 3.2).

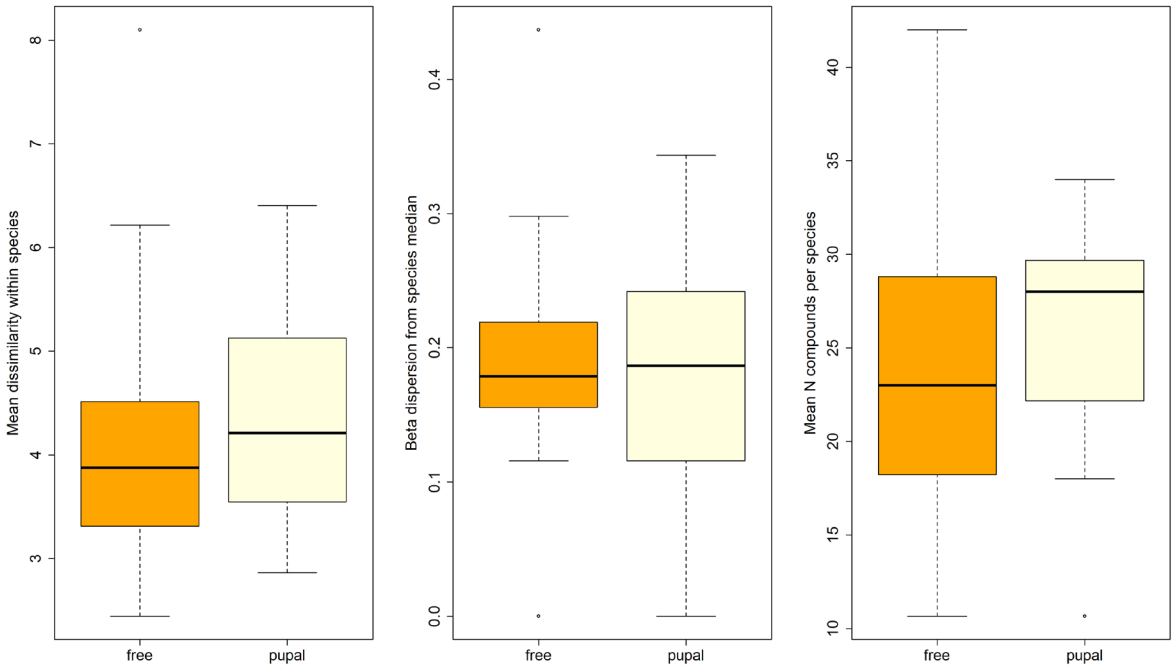
Figure 3.2. Phylogenetic distributions of the main early eluting compounds, including sulcatone, (E)-β-ocimene and 4-hydroxycyclopent-2-en-1-one. Compounds are ordered by RI, and colour represents the percentage of the blend represented by each compound. In order: octen-3-one, sulcatone, 2-methyl-2-heptene-6-ol, 4-hydroxycyclopent-2-en-1-one, phenylacetaldehyde, (E)-β-ocimene, 3-nonanone, O-guaiacol, cis-linalool oxide, (E)-epoxy-ocimene, benzylcyanide, 1-decen-3-one, (Z)-3-hexenylisobutyrate, 3-decanone, (Z)-3-hexenyl-2-methylbutanoate, hexyl-3-methylbutyrate, dihydroedulan II, 3-undecanone, 83-82-1306 (unknown, RI=1306), GC-EAD active compound (unknown), dihydro-β-ionol, unknown hexenyl ester, homovanillylalcohol, 127-55-1556 (unknown, RI=1556), unknown terpene, dibenzylamine.



Following this initial qualitative analysis, I calculated within-species dissimilarity, dispersion (distance from species median) and number of compounds and used these parameters to compare the free and pupal-mating clades, in order to validate the notion proposed in (Estrada *et al.*, 2011) that pupal-mating species would be evolving at a slower rate due to less competition between males and less sexual conflict. Under that assumption, conspecifics from free-mating species should also appear more diverse than those from pupal-mating species. Based on Welch's two-sample t-test, I found no significant difference between the *Heliconius* pupal- and the free-mating clades in terms of average within-species dissimilarity expressed as mean Euclidean distance between conspecifics (free-mating species = 4.1 ± 1.4; pupal-mating species = 4.4 ± 1.1; t = -0.36, df = 24.69, p = 0.71) or number of

different compounds (free-mating species = 23.6 ± 9.1; pupal-mating species = 25.4 ± 6.6; t = -0.59, df = 27.96, p = 0.56) (Fig. 3.5). Likewise, dispersion from species median was not significantly higher in free-mating species, with it corresponding to 0.19 ± 0.09 for free-mating species and 0.17 ± 0.11 for pupal- mating species (t = 0.48, df = 23.91, p = 0.64) (Fig. 3.5). Inclusion of *Eueides* and basal species (which are all free-mating) in the test alongside free-mating *Heliconius* does not alter the results of the test.

Figure 3.3. Comparison of mean within-species dissimilarity, dispersion from the species median, and mean number of compounds per species between pupal (13 species) and free-mating (16 species) *Heliconius.* None of these parameters are significantly different between these two groups of species.



These results suggest stability of the antiaphrodisiac blend within species, regardless of the mating strategy. This is further corroborated by the ADONIS test results, which show a very strong effect of species on pheromone composition (R^2 = 0.95, df=38, p<0.0001). When carried out on average species data, ADONIS also shows a significant contribution of clade on pheromone composition, indicating that closely related species tend to be similar (R^2 = 0.43, df=9, p<0.0001). The NMDS based on Euclidean distances accurately represents within-species and within-clade variation in pheromone composition (Fig. 3.3), and its shape seems to mainly be driven by the three most prominent early-eluting compounds. The plot also shows the differences in within-clade variation, with some clades (such as the *sara* and *cydno* clades) being more variable than others.

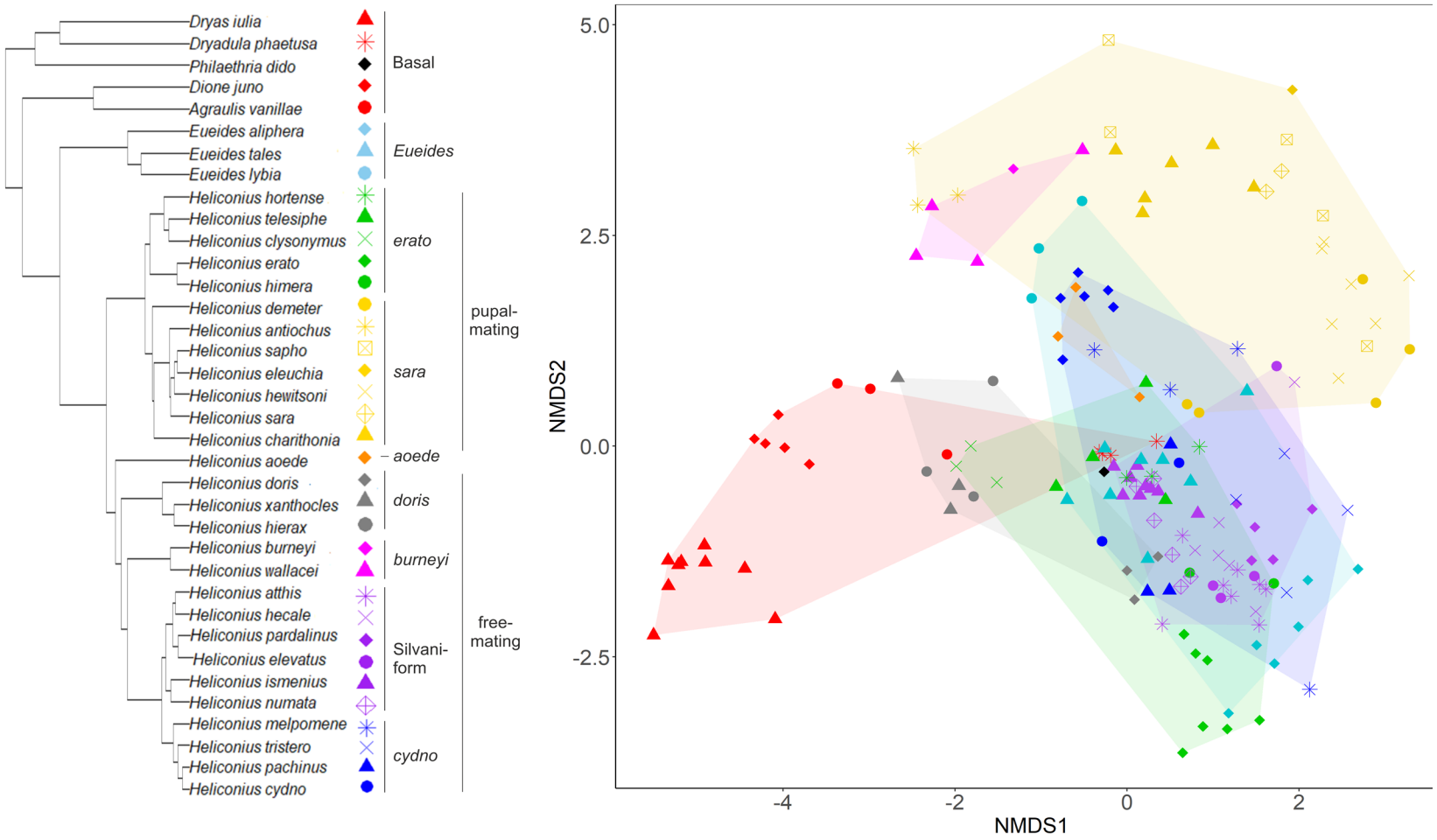


Figure 3.3 NMDS plot of all samples featured in the study, showing the dissimilarity between samples (expressed as Euclidean distance). This plot shows the first two axes out of 5 (k=5).The stress value for this k=5 NMDS model is 0.11, indicating a good fit to the data. Colors and hulls represent Heliconiini subclades, and each combination of shape and colour represents a single species as shown in the phylogeny.

I used a glm model to test for the effect of range overlap and branch length (divergence time in Myr) on the dissimilarity (Euclidean distance) of genital contents between sister species, to assess whether the former might drive antiaphrodisiac diversification in sympatric sister species pairs, a pattern first observed in the androconial extracts (Chapter 1). Branch length was used to control for phylogenetic effects. The model returned no significant effect of range overlap (Est.=-0.14, t=-0.99, p=0.35), but a significant effect of branch length (Est.=0.04, t=3.661, p=0.008) (Supp. Fig. 3.2). These results, compared those seen with androconial blends in Chapter 1, offer an interesting contrast between a trait that is likely to be acting as a pre-zygotic reproductive barrier (androconial pheromones) and one that is unlikely to take on that role. Closely related species seem to have similar genital blends regardless of sympatry, with branch length being the sole factor in this model affecting differentiation. This is reflected in the NMDS (Fig. 3.3), wherein closely related species tend to fall within the same region of the plot.

Phylogenetic analysis

Having carried out a qualitative analysis of genital blend composition, I then tested a few phylogenetic hypotheses. First, I tested how phylogenetically constrained the genital blend and its subsets (early and mid/late-eluting compounds) are, using the K_{mult} parameter, in order to discern whether the composition is more similar between closely related species (high phylogenetic signal, constrained composition, K≥1 or close to 1) or more similar between distantly related species (low signal, evolutionarily labile composition K closer to 0). K was also calculated for each compound individually, to investigate whether specific components are more constrained than others.

While in the qualitative analysis ADONIS revealed that the clade to which a species belongs makes a contribution to antiaphrodisiac composition, overall the blend carried little phylogenetic signal, and lower than that expected under a Brownian motion model of trait evolution (K=0.32, p<0.0001), concordantly with results from Estrada (2011). Early- eluting compounds are slightly more constrained (K=0.35, p<0.0001) than the mid/late-eluting matrix components (K=0.25, p=0.04) (Table 3.1A). When it comes to phylogenetic signal, single compounds behave differently from the blend as a whole, with several examples of blend components showing much higher phylogenetic signal and most of them showing no signal at all on their own as evidenced by the greatly variable K values shown in (Table 3.1B).

These results, given the low K values, show that the blend is not evolving in a simple Brownian mode and that antiaphrodisiac composition probably underwent one or more bursts of fast diversification. Results from model-fitting with mvMORPH further demonstrate this. The best fitting model, whether considering the full blend, the early-eluting compounds, or the mid/late-eluting matrix, is Brownian

motion followed by an early burst at the root of the advanced Heliconiini, corresponding with the *Eueides-Heliconius* split (Table 3.2).

A.

Compound group	Kmult	p-value
Full blend	0.2487	<0.0001
Early-eluting	0.2937	0.0128
Mid-eluting	0.3268	<0.0001

B.

Compound name or ID	RI	K	p-value
Sulcatone	993	0.6911	0.0057
(E)-epoxy-Ocimene	1135	0.4757	0.0413
Dihydro-β-Ionone	1441	0.3489	0.0455
Hexadecyl acetate	2018	0.4716	0.0350
83-55-2077	2077	0.2927	0.0321
Henicosane	2100	0.4326	0.0056
73-55-2166	2166	0.6173	0.0182
83-55-2286	2286	0.2936	0.0211
57-55-2338	2338	0.4267	0.0378
2-Acetoxyhexadecenoate	2346	0.5118	0.0298
Isoprenyl ester 3	2425	0.7544	0.0443
(Z)-13-Docosen-1-ol	2464	0.4819	0.0056
Sulcatyl ester 1	2649	2.0259	0.0011
Sulcatyl ester 3	2680	0.6979	0.0184
Heptacosane	2700	0.2754	0.0415
69-55-2719	2719	0.6695	0.0483
82-55-2786	2786	0.6831	0.0329
Sulcatyl ester 4	2883	0.4260	0.0420
Nonacosane	2900	0.3615	0.0090
Isoprenyl ester 10	2909	0.4758	0.0407
340-339-2972	2972	0.7205	0.0413
55-81-2999	2999	0.5498	0.0436

*Table 3.1*A. K_{mult} results for the whole blend and the two subsets. B. List of all single compounds that show significant K values. K values close to 0 indicate low phylogenetic signal (the compound's relative amount is more similar between distantly related species than between closely related species, meaning the amount is evolutionarily labile); K=1 or close to 1 indicates the compound is evolving under Brownian expectations, meaning it is not evolutionarily labile; K>1 indicates the compound's amount is more phylogenetically constrained than expected under Brownian assumptions. Colors reflect this, with red indicating low signal and green indicating high signal. Where the compound name is unknown, the compound is indicated with a three-number identifier where the first two numbers are the two most abundant m/z fragments as reported by the MS and the third number is the retention index (RI).

The MOTMOT tm1 model gave different results: it did detect increases in evolutionary rate at the node between the free-mating *H. tristero* and *H. cydno/pachinus* (AICc=612) and at the base of the pupal-mating clade (AICc=605), but neither was a significantly better fit than the base shiftless Brownian model (AICc=626), based on an AICc cutoff of 20.23 (Supp. Table 3.1). Therefore, unlike

Estrada et al. (2011), neither mvMORPH nor MOTMOT found any significant rate shifts corresponding with shifts in mating strategies.

Past results on *H. melpomene* (Schulz *et al.*, 2008) suggested slightly different functions for early-eluting compounds and for the matrix (mid/late-eluting compounds): early-eluting compounds would carry the strongest antiaphrodisiac signal, with the matrix (mid/late-eluting compounds) having a weaker behavioural impact and mainly serving the purpose of modulating the effect of the formers. Based on this, the two categories of compounds might evolve at different rates, with the formers being more conserved and therefore phylogenetically constrained, and the latters being more labile, potentially communicating finer information such as species or male identity. Based on this, I used mvMORPH to fit the same models to two separate datasets: early eluting (RI <1800) and mid/late-eluting (RI>1800). However, in both cases, a shift was only detected at the *Eueides/Heliconius* node. For mid-eluting compounds, the best fitting model was Brownian motion followed by a burst at the *Eueides/Heliconius* node, as was the case for the full blend (Table 3.2), whereas for early-eluting compounds the best fit was similarly given by a Brownian motion model with a simple rate shift at *Eueides/Heliconius*. When using the MOTMOT tm1 model, the best fitting model for mid/late-eluting compounds is basic shiftless Brownian motion, whereas on early-eluting models the best fitting model includes a ~x5 increase in rate at the root of *Heliconius* (Supp. Table 3.1). Thus, even when considering different sets of compounds, no shift concurrent with mating strategy was detected.

Even though the best fitting models for evolutionary mode never included a significant rate shift concurrent with shifts in mating strategies, I used the r(geomorph) function compare.evol.rates to quantify the difference in evolutionary rates between the two mating strategies within *Heliconius* (non-*Heliconius* species were excluded from these calculations). Rate comparisons between the free- and pupal-mating clades yielded similar results to those obtained from mvMORPH and MOTMOT: I detected no significant difference in rate between the two lineages in the full blend or either of the two subsets (observed rate-ratio = 1.2472, 95% CI=[1.01,1.47], p=0.24), with the evolutionary rate σ being very similar between them (δ=0.05 for pupal-mating species and σ=0.04 for free-mating species).

Model				
Full blend	AIC	diff	wi	AICw
BM followed by EB at *Heliconius/Eueides* lineage	**657.6971**	**0**	**1.0000**	**0.9859**
Simple Brownian motion	667.5396	9.84253	0.0073	0.0072
EB followed by BM at pupal-mating lineage	669.5637	11.86666	0.0026	0.0026
BM followed by EB at free-mating lineage	669.83	12.13295	0.0023	0.0023
Early-Burst	670.1047	12.40761	0.0020	0.0020
BM with rate shift at pupal-mating lineage	680.5433	22.84619	0.0000	0.0000
BM with rate shift at free-mating lineage	682.5608	24.86371	0.0000	0.0000
Early-eluting	AIC	diff	wi	AICw
BM with shift at *Heliconius/Eueides* lineage	**499.8808**	**0**	**1.0000**	**0.9999**
BM followed by EB at *Heliconius/Eueides* lineage	518.9351	19.05435	0.0001	0.0001
BM with shifts at *Heliconius/Eueides* and free-mating lineages	524.2209	24.34014	0.0000	0.0000
Simple Brownian motion	538.6387	38.75794	0.0000	0.0000
BM followed by EB at free-mating lineage	538.8686	38.98786	0.0000	0.0000
EB followed by BM at pupal-mating lineage	541.0461	41.16533	0.0000	0.0000
Early-Burst	541.1329	41.25212	0.0000	0.0000
BM with shift at pupal-mating lineage	556.9248	57.04405	0.0000	0.0000
BM with shift at free-mating lineage	557.0608	57.18006	0.0000	0.0000
Mid-eluting	AIC	diff	wi	AICw
BM followed by EB at *Heliconius/Eueides* lineage	**634.4562**	**0**	**1.0000**	**0.9954**
Simple Brownian motion	646.9366	12.48038	0.0019	0.0019
BM followed by EB at free-mating lineage	647.2961	12.83984	0.0016	0.0016
EB to BM at pupal-mating lineage	649.5013	15.0451	0.0005	0.0005
Early-Burst	649.5017	15.04546	0.0005	0.0005
BM with shift at free-mating lineage	665.111	30.65476	0.0000	0.0000
BM with shift at pupal-mating lineage	666.642	32.18579	0.0000	0.0000

Table 3.2. Results of mvMORPH model fitting ranked by best fit to the full blend and the two subsets of compounds. AIC is the corrected Akaike information criterion, used to identify the best-fitting model, and diff is the difference between AIC scores between the models; wi is the model weight, and AICw shows the fit of the model relative to all others. Lower values of AIC indicate a better fit of the model to the data, and values of wi and AICw very close to 1 indicate a very good fit relative to other tested models. Not listed are any models that could not be reliably fit to the data, which produce disproportionately low AIC values and interfere with wi and AICw calculations.

Restricting the analysis to just the species included in Estrada et al. (2011) also gives results discordant with the 2011 paper: with no significant difference in diversification rate (observed rate-ratio = 1.048, 95% CI=[1.02, 1.86], p=0.87) between the free-mating clade (σ=0.03) and the pupal-mating clade (σ=0.04).

Phylogenetically corrected correlations were found between the relative amounts of some groups of compounds, albeit most compounds in the datasets appear not significantly correlated to each other (Fig. 3.4). Some of these compound clusters are based on the compound class: alkanes form some scattered groups based on their weight, and several unidentified alcohols, aldehydes and esters also

show strong correlations to each other. The prominent early-eluting compound (E)-β-ocimene shows remarkably few correlations, seemingly only being correlated to the alcohol (Z)-13-docosenol and to two unidentified compounds (Fig. 3.4). The other widespread early eluting compounds also display strong associations with matrix components. 4-hydroxycyclopent-2-en-1-one is associated with phytol and some unidentified matrix components, including plant derivatives such as homovanillylalcohol and an unknown terpene (Fig. 3.4). The relative amount of sulcatone (6-methyl-5-hepten-2-one) is correlated with several other compounds, but most strikingly with three specific matrix components, the sulcatyl esters, which are also its derivatives, with which it co-occurs both in the basal species and in *H. doris* (Fig. 3.4). Sulcatone shows relatively strong phylogenetic signal and the same is often true for its sulcatyl esters (Table 3.1B), possibly due to being strongly conserved among basal species in spite of their less recent divergence. Sulcatone however also correlates to many other types of matrix components including the esters hexadecyl acetate, 2-acetoxyhexadecanoate, 1-hexadecenyl acetate, heptadecyl acetate and octadecyl acetate, as well as other early-eluting compounds such as 2-methyl-2-hepten-6-ol, dihydroedulan II; dihydro-β-ionol and dihydro-β-ionone (Fig. 3.4). Octen-3-one, meanwhile, shows a strong correlation with a similar compound, 1-decen-3-one, and with the alcohol 1-octadecanol (Fig. 3.4).

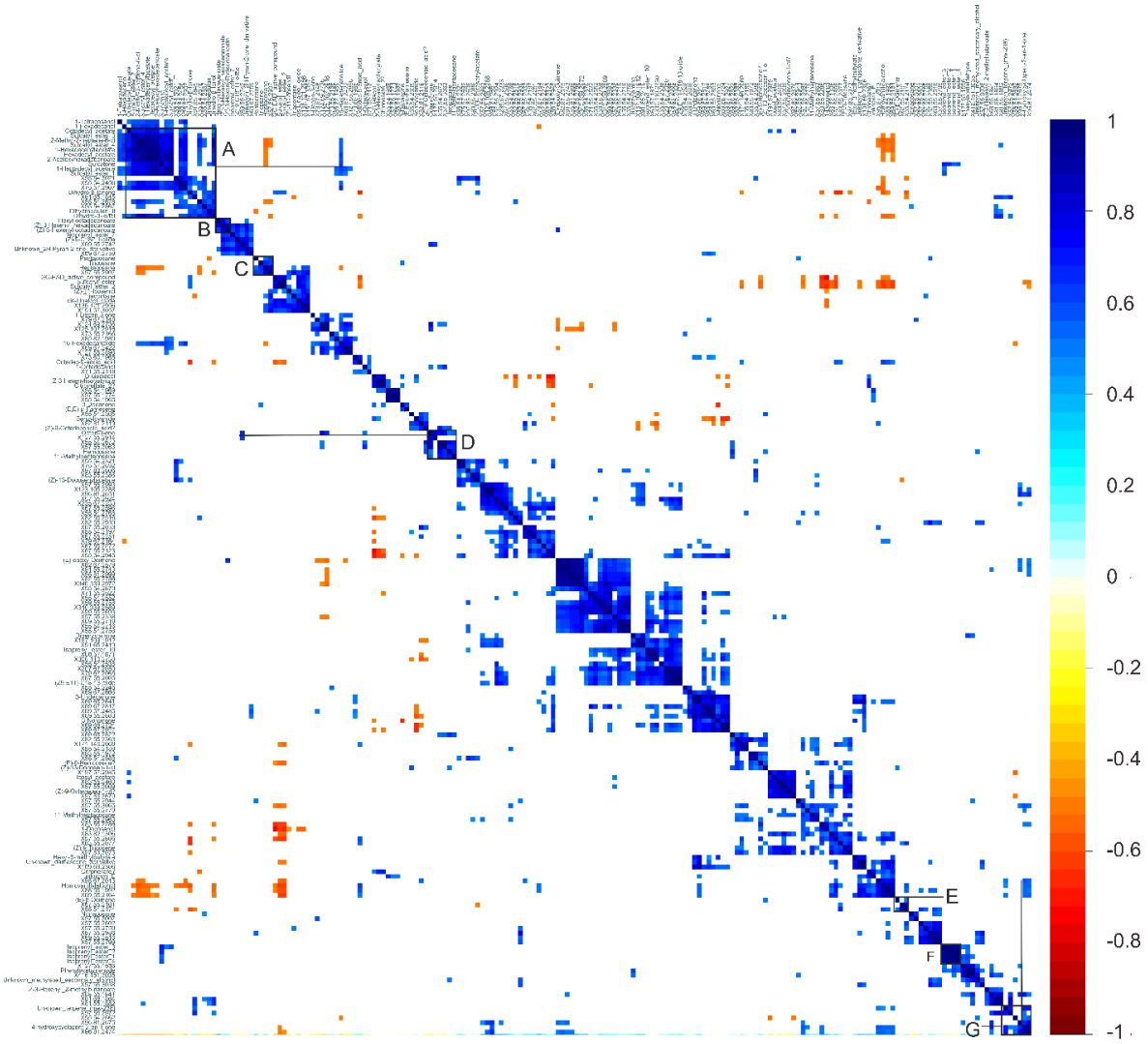


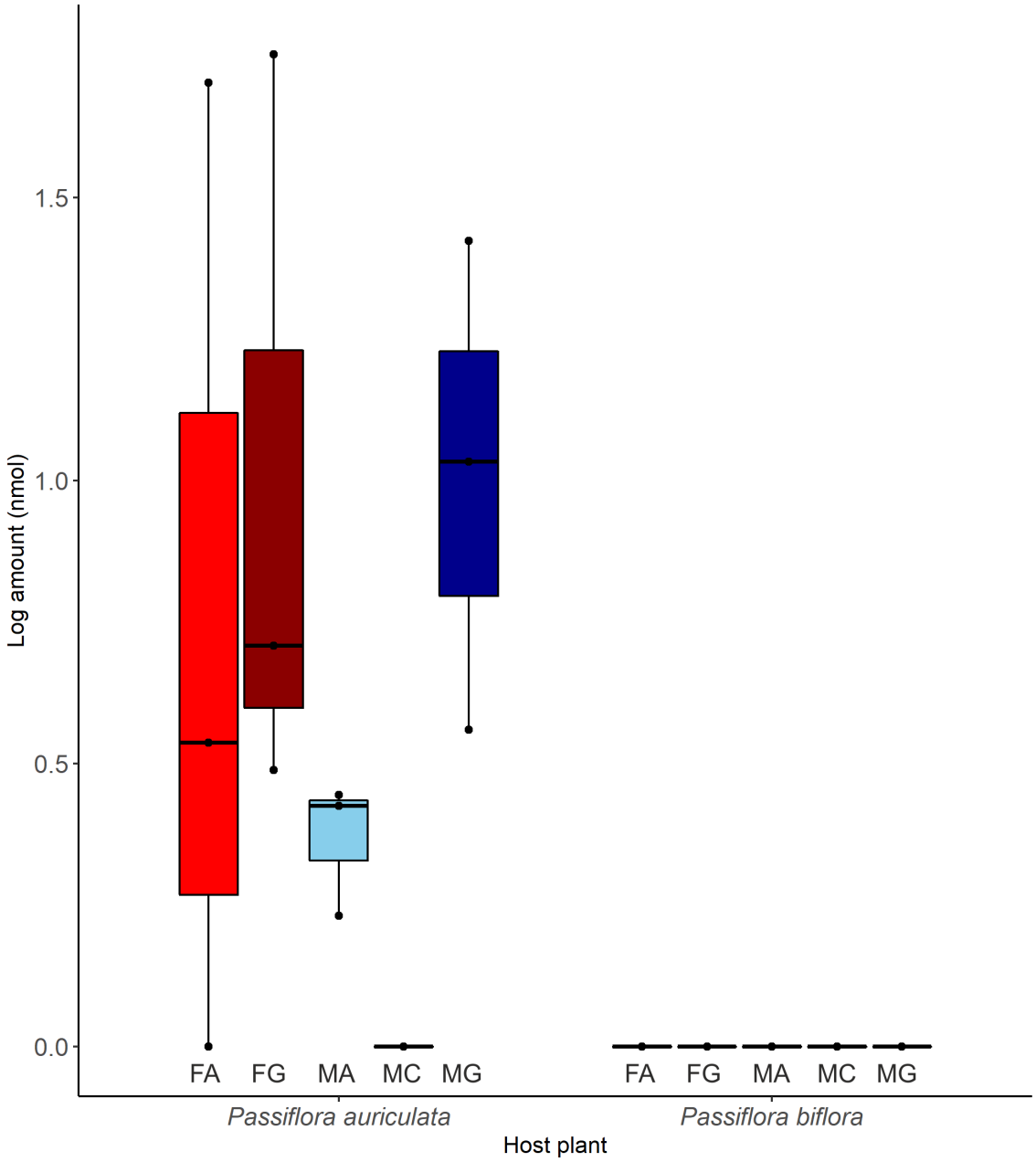
Figure 3.4. Pairwise correlation matrix for all compounds in the dataset appearing more than 3 times across all species averages. Colour represents the Pearson correlation coefficient, and blank tiles being non-significant correlations (p-values corrected with the Benjamini-Hochsberg procedure). Squares indicate some of the clusters of known compounds, including 4 clusters for the main early-eluting compounds (sulcatone, (E)-β-ocimene, 4-hydroxycyclopent-2-en-1-one and octen-3-one). Lines, where present, indicate compounds that correlate to one of the 4 main early-eluting compounds but fall outside of the cluster. A) Sulcatone clusters, including sulcatyl esters, fatty alcohol esters, dihydroedulan II, dihydro-β-ionone, and dihydro-β-ionol and 2-methyl-2-hepten-6-ol; B) cluster of hexyl and (Z)-3-hexenyl esters; C) cluster of some long chain (23C) alkanes; D) octen-3-one cluster, including several unidentified compounds as well as 1-decen-3-one and 1-octadecanol; E) (E)-β-ocimene cluster, including 2 unidentified compounds and (Z)-13-docosenol; F) cluster of isoprenyl esters; G) cluster for 4-hydroxycyclopent-2-en-1-one, including some plant derivatives (an unidentified terpene and homovanillylalcohol) and other largely unidentified alcohols and aldehydes.

Early-eluting compound occurrence in *Heliconius sara*

In order to explore the potential role of 4-hydroxycyclopent-2-en-1-one as an antiaphrodisiac in *Heliconius sara,* I first compared the levels of this compound of interest in freshly emerged males, older (sexually mature) males, virgin (freshly emerged) females and mated females. If the compound

is an antiaphrodisiac, it is expected to be present in sexually mature males and in mated females at significantly higher levels than in virgin females. However, 4-hydroxycyclopent-2-en-1-one was detected in none of these samples. These initial results informed a subsequent host plant experiment. *H. sara* is known to prefer *Passiflora auriculata* as a host plant in the wild (Brown, 1981), whereas *H. sara* from the University of Sheffield stocks were reared on *P. biflora.* I compared the genital contents of 6 individuals that were reared on *P. auriculata* with those of 6 individuals that were reared on *P. biflora* (3 males and 3 females in both cases). The compound was completely absent from individuals fed on *P. biflora,* but it was detected in individuals that fed on *P. auriculata* (Fig. 3.5). In these individuals, 4-hydroxycyclopent-2-en-1-one was found in the genitals of both males and females, with no significant difference in amount between the two (t-test, t=-0.16, df=3, p=0.88). Furthermore, the compound of interest was found in the androconia of both males and females, though it was never found in control wing samples from male individuals, indicating that while it may have no purpose as an antiaphrodisiac, it is not a ubiquitous compound that is equally incorporated from the host plant into all body parts of the organism.

Figure 3.5. Logged amount of 4-hydroxycyclopent-2-en-1-one across the different male and female tissue types in *Heliconius sara* fed on two different host plants (*Passiflora biflora* and *Passiflora auriculate*) demonstrating how the presence of this compounds is strongly dependent on larval food. The tissue types are female "androconia" (FA), female genitals (FG), male androconia (MA), male control (MC) and male genitals (MG).



Discussion

This study generated a vast amount of new data on Heliconiini chemical ecology. I successfully performed a non-targeted GCMS analysis on the genital extracts of 36 Heliconiini species, 24 of which had never been analyzed for their antiaphrodisiacs and genital volatiles in the past. The non-targeted approach, relatively uncommon in the field of insect chemical ecology, was necessary as a large part of the compounds in the dataset had not been previously documented in *Heliconius.* The Heliconiini genital blends were revealed to be extremely complex and varied, with this striking diversity potentially resulting from a burst in diversification that took place at the base of the *Eueides* and *Heliconius* split. In spite of this detected burst, every other "basal" species present in the study (*Dryas iulia, Dione juno, Dryadula phaetusa, Agraulis vanillae* and *Philaethria dido*) also showed rich and complex genital blends. Most of these species had one or a few prominent early eluting compounds, a pattern previously observed in *Heliconius* antiaphrodisiacs (Schulz *et al.*, 2008; Estrada *et al.*, 2011) and in particular in *H. melpomene,* where the early-eluting compound has a powerful behavioural effect in conjunction with the matrix of heavier, mid/late-eluting compounds (Schulz *et al.*, 2008). This might imply that all or most species wherein this pattern is observed may utilize the early eluting volatile as the main antiaphrodisiac component, although I found no evidence of this being the case on at least one early-eluting volatile, 4-hydroxycyclopent-2-en-1-one.

Unlike previously documented (Estrada *et al.*, 2011), I was not able to detect any effect of mating strategy on pheromone macroevolution. In both the pupal and the free-mating lineage, the extracts were remarkable species-specific, and showed no more intraspecific variation in the free-mating lineage than in the pupal-mating lineage. Likewise, no shift in evolutionary rate or mode was detected concurrently with either mating strategy, providing no evidence that this aspect of *Heliconius* biology has affected their antiaphrodisiac evolution.

My results confirm past findings (Schulz *et al.*, 2008; Estrada *et al.*, 2011) that Heliconiini antiaphrodisiac blends are very complex and strikingly variable among species. PyMassSpec, here used for the first time to analyze pheromone samples, turned out to be a valuable tool for expediting the data deconvolution process, in conjunction with appropriate retention index data produced via the standard alkane ladders ran alongside the samples.

Much like the chemical blends produced by male *Heliconius* in their wing androconia (Vanjari *et al.*, 2015; Mann *et al.*, 2017, Cama, unpublished; Chapter 1), genital extracts are comprised of several different classes of compounds, including alcohols, aldehydes, ketones, terpenes and esters of varying complexity, and the near-ubiquitous fatty acids, alkanes and alkenes, and like the androconial blends,

every blend is a mixture of plant-derived and fatty acid-derived components, though some of the formers may actually be produced endogenously as opposed to being transformed from host plant compounds, as seen with (E)-β-ocimene in *H. melpomene* (Darragh, Orteu, *et al.*, 2019).

One notable difference between genital and androconial blends is the presence of a great variety of fatty acid esters in the basal genera *Dryas, Dryadula, Dione* and *Agraulis,* whereas these species' androconial blends contain almost no fatty acid derivatives. It is important to note, however, that the antiaphrodisiac function of most of these compounds remains unconfirmed, with the notable exception of (E)-β-ocimene (Schulz *et al.*, 2008; Darragh, Orteu, *et al.*, 2019). Therefore, most of them may either be structural in nature, not behaviourally active, or not involved in sexual signalling at all.

Potential purpose of early eluting compounds

ADONIS results have shown that the genital blends are largely species-specific, even though most of this variability lies in the matrix components. When it comes to early-eluting compounds, most species tend to have one of 4 four compounds: (E)-β-ocimene, 4-hydroxycyclopent-2-en-1-one, sulcatone and octen-3-one. Despite there being little variation in the types of prominent early-eluting compounds across species, the amounts of these compounds relative to the rest of the blend are very variable, especially in for (E)-β-ocimene which ranges from a mean of 1.8% of the total blend in *H. hortense* to 83.2% in *Dryadula phaetusa* (Fig. 3.2). It is currently unclear whether there are any patterns in the relative amounts of these compounds and what they might be driven by.

The pattern in the chromatographic data of one/a few early-eluting compounds followed by a matrix was initially first observed in *H. melpomene* (Schulz *et al.*, 2008)*,* then later described in more *Heliconius* species (Estrada *et al.*, 2011) and here I have found it to be common across the entire Heliconiini tribe including non-*Heliconius* genera. All basal species as well as *Eueides aliphera* show a similar pattern. In *H. melpomene,* (E)-β-ocimene, a member of the terpene class of compounds, is the only known behaviourally-active blend component (Schulz *et al.*, 2008), and it is produced by a specific and newly discovered terpene synthase not related to those found in other terpene-producing organisms (Davies, 2008)organisms, be they plants or insects (Darragh, Orteu, *et al.*, 2019). This compound has turned out to be very widespread among the Heliconiini: it is present in most *Heliconius* clades analyzed thus far (*erato, cydno* and silvaniform) and in all of these it is produced by at least one species in very large amounts, and it is also found in the basal species *D. phaetusa* and *P. dido* where it makes up a large proportion of the blend. Overall, this compound is a very widespread pheromone among insects, found in other Lepidoptera (Honda, 1981; Hayashi *et al.*, 1985), in Hymenoptera (Keegans *et al.*, 1993; Eltz *et al.*, 2006), and in Coleoptera (Giglio *et al.*, 2009; Sun *et al.*, 2010), among many others. In all of these examples, (E)-β-ocimene has taken on a variety of different functions in

chemical communication, from attraction to defense, and the life stage in which it is produced also varies.

Given the established role of (E)-β-ocimene, one might speculate that the other three main early-eluting compounds, sulcatone, octen-3-one and 4-hydroxycyclopent-2-en-1-one, could have similar roles in the species where they occur. Of the three, sulcatone is better known for its role as a pheromone in many species. Unlike (E)-β-ocimene, this compound is found almost exclusively in some basal species within the Heliconiini (*Agraulis vanillae, Dione juno* and *Dryas iulia*), with the exception being *H. doris.* There is one past report of its occurrence in *A. vanillae,* where it was presented as a predator repellent (Ross *et al.*, 2001), however this does not exclude its potential function as an antiaphrodisiac pheromone. It is also a known alarm pheromone in some species of ants (Tomalski *et al.*, 1987) and a defense pheromone in a genus of parasitoid wasps (Hübner *et al.*, 2002), and it is even found in some vertebrates such as otters (*Lutra lutra*) and *Hypsiboas* tree frogs (Brunetti *et al.*, 2015). On the other hand, octen-3-one is potentially utilized as an alarm pheromones in fungus-growing ants (Norman *et al.*, 2017).

Meanwhile, 4-hydroxycyclopent-2-en-1-one is known for its role in Heliconiini toxicity to predators (Pinheiro de Castro *et al.*, 2019), but its putative function as a pheromone is almost entirely unexplored, with no publications pertaining to its role as a behaviourally active compound in any species. It is found in the androconial and genital extracts of Panamanian field-collected wild *Heliconius sara* males in significantly higher amounts than in those of females from the same region (Chapter 1), suggesting a potential function in courtship and mating. However, captive specimens from the University of Sheffield utilized in our host- plant experiment revealed no such sex-based difference in either the androconial or the genital secretions.

There are a few possible reasons for this discrepancy. Geographic intraspecific differences in pheromone composition are very common and thoroughly documented, and with them come differential female reactions to the pheromone blend components (Rollmann, Houck and Feldhoff, 2000; McElfresh and Millar, 2001; El-Sayed *et al.*, 2003; Palacio Cortés *et al.*, 2010). Thus, it is possible that the Panamanian *H. sara* population utilizes 4-hydroxycyclopent-2-en-1-one as a pheromone component, while geographically distant populations may not. Pheromone production is also affected by abiotic factors such as light, temperature and resource availability in several species, which would lead to noticeable differences between wild and captive populations (Vanderwel, 1994; Parker and Mason, 2009; South *et al.*, 2011; Bontonou, Denis and Wicker-Thomas, 2013; Eller and Palmquist, 2014; Rudie, 2015; Darragh, Byers, *et al.*, 2019).

Most strikingly, I found that consumption of *Passiflora auriculata,* the preferred host plant for *H. sara,* results in the production of 4-hydroxycyclopent-2-en-1-one in captive individuals. The literature is rife with examples of diet-affected pheromones (e.g. Löfstedt *et al.*, 1989; Kopena *et al.*, 2011; Fedina *et al.*, 2012; Liedo *et al.*, 2013; Henneken *et al.*, 2015; reviewed in Henneken *et al.*, 2017). Overall, our results highlight the importance of considering the origin of experimental specimens when carrying out pheromone-related research in butterflies. Regardless of the source of the difference between Panamanian and captive *H. sara*, this makes 4-hydroxycyclopent-2-en-1-one an unlikely candidate for *H. sara* antiaphrodisiacs, meaning that early-eluting compounds in the Heliconiini may not always have a behavioural function as strong as that of (E)-β-ocimene in *H. melpomene*.

Matrix compounds: antiaphrodisiac modulation?

Even if the matrix components may not be behaviourally active, it has been suggested that they may be involved in the modulation of early-eluting compounds evaporation times, thus regulating the duration of the antiaphrodisiac effect (Schulz *et al.*, 2008). It is also possible that the matrix components may be synergizing with the early-eluting compounds to modify or enhance the behavioural response of other males, even if said synergy may not be given so much by chemical reactions between the components, but rather by an alteration of the male's perception of the signal. At least one example of this is also seen in the wasp *Eriborus terebrans,* where a behaviourally- active compound synergizes with a structural component of the cuticle to elicit a stronger response (Shu and Jones, 1993). It is also not uncommon in insects for plant volatiles to enhance the effect of pheromones (Reddy and Guerrero, 2004), especially among Coleoptera (Collignon *et al.*, 2016; Ju *et al.*, 2017); while in most known cases the greater response is produced by the combination of compounds from two distinct organisms (the insect-emitted pheromones and the host plant-emitted volatiles), plant-derived components of the *Heliconius* genital blend might modify the effect of other *Heliconius*-produced volatiles. Unfortunately, not much is known about how male Heliconiini perceive the antiaphrodisiac signal, so the function of all these compounds remains difficult to decode for now. Another possibility, given the little variation in early eluting compounds even among close relatives, is that the matrix may carry a signal of species identity, which may explain its greater variability between different species when compared to early-eluters. The matrix compounds may also fulfil more than one of these purposes simultaneously.

Regardless of what the purpose of the matrix is, past experiments have shown that its components are necessary for optimal antiaphrodisiac activity of more volatile compounds to take place (Schulz *et al.*, 2008). The strong and phylogenetically independent correlations between early-eluters and clusters of matrix components also hints at a synergistic purpose of the matrix components, and/or

implies shared genetic or biosynthetic control. The latter seems especially obvious in the case of sulcatone and its "matrix companions", the sulcatyl esters, which are also its derivatives.

Antiaphrodisiac evolution in the face of a shift in mating strategy

Methods for phylogenetic analysis of multivariate traits, not yet available at the time of the last *Heliconius* antiaphrodisiac macroevolution study (Estrada, 2011), had a great impact on the way one may interrogate chemical ecology data. This, in combination with the addition of 24 new species not previously featured in the study, produced very different results from those obtained in the past. In contrast with past findings from Estrada (2011), I did not detect a shift in evolutionary rate at in either the free-mating or the pupal- mating lineage, with or without *a-priori* input. The previously reported higher rate of evolution in the free-mating clade may have been a consequence of a sparse sampling of species reported in Estrada et al. (2011), although I also attempted to utilize multivariate phylogenetic model-fitting methods on the same subset of species featured in Estrada (2011) and did not find evolutionary shifts with either mating strategy.

Overall, it seems like the shift in mating strategy has not had any impact on antiaphrodisiac evolution. However, rate shifts may have happened at a finer level: the *sara* clade is best known to engage in pupal -mating, and this behaviour has been studied in detail in *H. charithonia* and *H. hewitsonii* in particular (Estrada and Jiggins, 2008; Estrada *et al.*, 2010; Mendoza-Cuenca and Macías-Ordóñez, 2010; Thurman *et al.*, 2018). However, the *erato* clade has a very different ecology from the *sara* clade (Thurman *et al.*, 2018), so it is possible that any pheromone-related adaptations driven by pupal mating may only be found in the latter, albeit the *sara* clade genital extracts actually appear more varied than the *erato* ones in this study (Fig. 3.3). Nonetheless, I found no evidence for shifts at any of these levels, and given the small number of species within each clade, it might be difficult to reliably detect them (Adams and Collyer, 2018). It is also important to note that just because no effect of mating strategy has been found on the evolution of antiaphrodisiac composition, it does not mean that the antiaphrodisiac system has remained entirely unaffected. Selection may have acted on other aspects of this signalling system, such as male perception or investment, and regulation of antiaphrodisiac production.

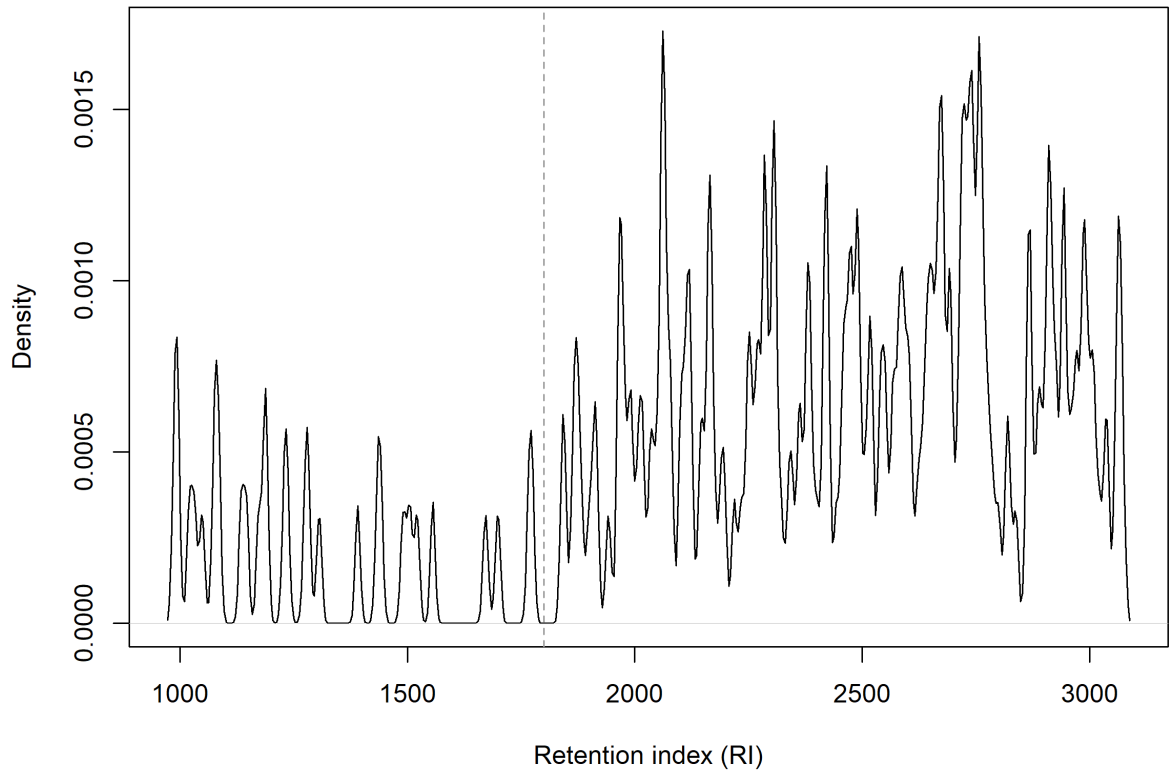
The addition to the analysis of more species did not lead to an increase in detectable phylogenetic signal, and the low values for K_{mult} indicate that genital blends are not very phylogenetically constrained (Adams, 2014), in concordance with past results, albeit even though singular individual components of the blend occasionally display very high phylogenetic signal. Among these are several of the early-eluting volatiles, compounds they are strongly correlated with, and structural alkanes. When analyzed as a group, early-eluting volatiles are less labile than the matrix, as they tend to be

conserved within groups of closely related species. This fast diversification has coincided with the beginning of the advanced Heliconiini split, as is often seen with trait evolution in evolutionary radiations (Losos, 2010; Yoder *et al.*, 2010).

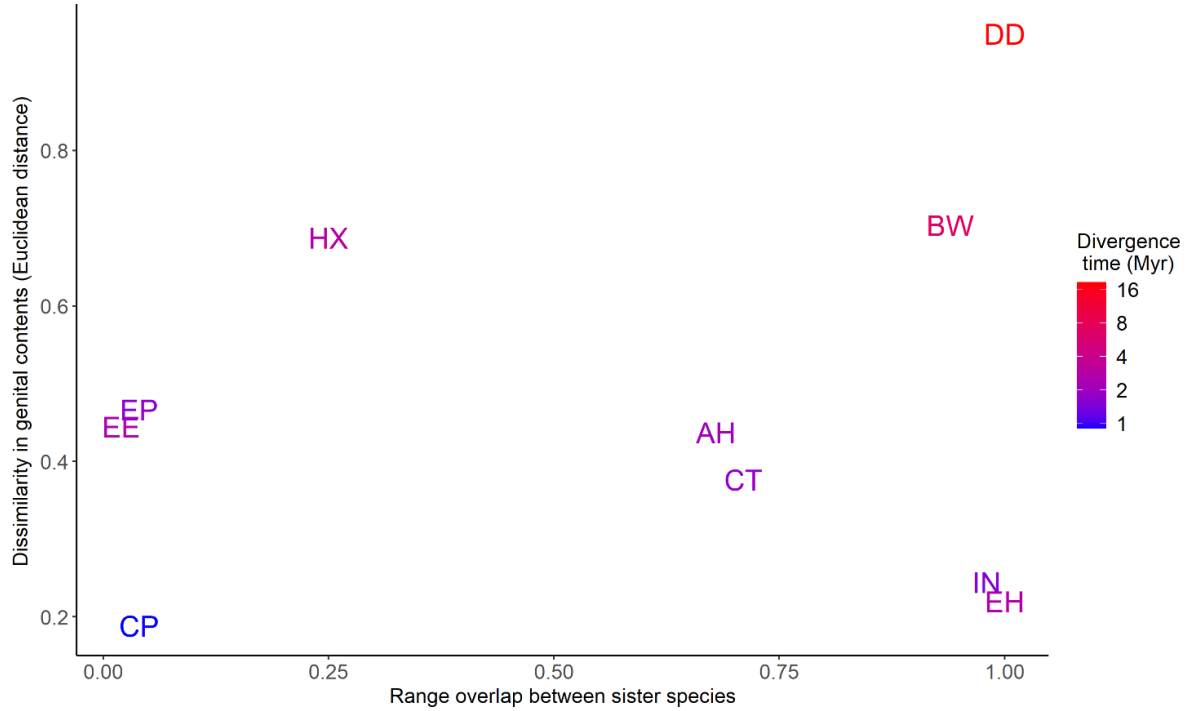
In male androconial blends, the opportunity to exploit different niches of pheromone composition was seemingly given by the upregulation of a fatty-acid metabolic pathway in *Eueides* and *Heliconius,* but it is unknown what opportunity led to fast diversification of antiaphrodisiac blends. One might hypothesize that the activation of the necessary biosynthetic pathways for the production of (E)-β-ocimene and 4-hydroxycyclopent-2-en-1-one may lead to the production of a wide variety of by-products, but the former is also found in basal genera, and the number of compounds that correlate strongly with either of them is very small compared to the large variety of chemicals compounds seen in the full dataset. As the matrix is where more inter-specific variation lies, I need to understand more about its composition and function before I can speculate as to why it evolved at such a fast rate.

Conclusions

The sheer diversity in chemical composition I have been able to observe in the genital extracts of our study species makes it difficult to find any broad patterns in the ways these pheromones have evolved. I have found no significant effect for a shift in mating strategy on pheromone evolution, unlike what had previously been suggested. While the "few early-eluting compounds followed by a rich matrix" chromatographic characteristic arrangement is very common in the Heliconiini, based on our investigation of 4-hydroxycyclopent-2-en-1-one and comparing our results from those obtained from (E)-β-ocimene (Schulz et al., 2008), it has become apparent that not all species are using similar mechanisms to deliver the antiaphrodisiac signal, with early-eluting compounds not being necessarily involved in male-specific chemical signalling. Additionally, while the presence/amounts of some components are strongly correlated, there does not seem to be a strong relationship between matrix composition and the identity of the early-eluting compounds, meaning different subsets of the blend components have probably evolved independently of one another, making for a strikingly complex evolutionary history. Importantly, I was able to demonstrate that host plant was a determining factor for the chemical signature of *H. sara.* This finding highlights how the use of captive individuals as opposed to wild ones, while surely having many merits (e.g. known age and mating history of each individual), requires careful consideration, as the results may not accurately reflect what is happening in the wild. Overall, the *Heliconius* antiaphrodisiacs, and antiaphrodisiacs as a chemical signalling method in general, remain poorly understood, but we are getting closer to potentially learning more about the way they function and have evolved in these non-model species.



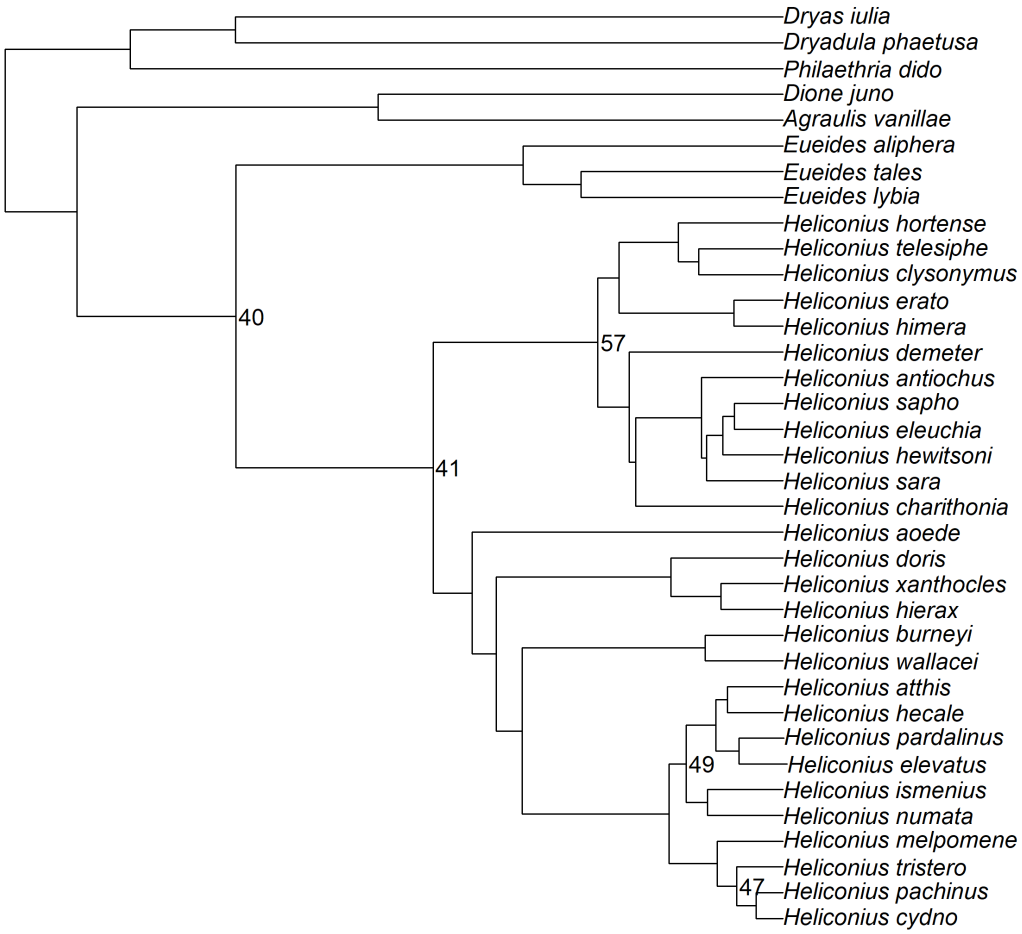
Supp. Figure 3.1 Plot of the density of compounds detected within 5-unit retention index windows. Lower retention indices indicate higher volatility. The dotted line marks RI=1800: lower RI compounds are classified as early-eluting, higher RI compounds as mid/late eluting.



Supp. Figure 3.2 Dissimilarity in genital contents between sister species, plotted against range overlap (Rosser *et. al,* 2015). Colors represent branch length (time since divergence in Myr) (Kozak *et. al*, 2015). The gradient in colors seen in this plot shows that branch length is the only factor affecting sister species dissimilarity, whereas there is no effect of range overlap. The pairs are (in descending order of branch length): *Dryas iulia-Dryadula phaetusa* (DD), *H. burneyi-H. wallacei* (BW), *H. hierax-H. xanthocles* (HX), *H. erato-H. himera* (EH), *E. lybia-E.tales* (EE), *H. atthis-H. hecale* (AH), *H. clysonimus-H. telesiphe* (CT), *H. elevatus-H. pardalinus* (EP), *H. ismenius-H. numata* (IN) and *H. cydno-H. pachinus* (CP).

Supp. Table 3.1. Results from model fitting with the r(MOTMOT) "tm1" (TraitMedusa1) model (Puttick *et al.*, 2020), that searches for rate shifts along all nodes in the phylogeny, on the full blend, on early-eluting compounds (RI<1800) and on mid-to-late-eluting compounds (RI>1800). The rates are reported in the "rate 1" and "rate 2" columns (no model found more than two potential rate-shifts). AICs is the corrected Akaike information criterion, used to identify the best-fitting model. Best-fitting model for each subset of the data are shown in bold, based on the AICc cutoff obtained from the CalcCutOff function in r(MOTMOT). The shift locations are indicated with numbers corresponding to their nodes, shown in the phylogenetic tree below.

Full blend	node	AICc (cutoff=20.227)	rate 1	rate 2
BM	**0**	**626.3951**	**NA**	**NA**
shift.1	47	611.873	6.767157	NA
shift.2	57	605.972	2.17392	9.375591
Early-eluting				
BM	0	511.5246	NA	NA
shift.1	**41**	**491.2114**	**5.724708**	**NA**
shift.2	49	475.9752	1.06762	9.387719
Mid-eluting				
BM	**0**	**605.7922**	**NA**	**NA**
shift.1	47	594.7552	5.31543	NA
shift.2	40	588.761	3.515022	16.8685


Dryas iulia
Dryadula phaetusa
Philaethria dido
Dione juno
Agraulis vanillae
Eueides aliphera
Eueides tales
Eueides lybia
Heliconius hortense
Heliconius telesiphe
Heliconius clysonymus
Heliconius erato
Heliconius himera
Heliconius demeter
Heliconius antiochus
Heliconius sapho
Heliconius eleuchia
Heliconius hewitsoni
Heliconius sara
Heliconius charithonia
Heliconius aoede
Heliconius doris
Heliconius xanthocles
Heliconius hierax
Heliconius burneyi
Heliconius wallacei
Heliconius atthis
Heliconius hecale
Heliconius pardalinus
Heliconius elevatus
Heliconius ismenius
Heliconius numata
Heliconius melpomene
Heliconius tristero
Heliconius pachinus
Heliconius cydno
40
57
41
49
47

Chapter 4: Pheromone production is concurrent with sexual maturity in *Heliconius* butterflies

Abstract

Among the many functions they can have in living organisms, pheromones are essential to animal reproduction as they allow individuals to find and choose a mate. As their function can be related to mating, production of pheromones in many organisms is tightly controlled and dictated by several external and internal conditions such as time of the day, season, proximity of conspecific individuals, and age. Although male sex pheromones and antiaphrodisiacs have been studied in some *Heliconius* butterfly species, little is known about the factors that affect their production and release in these relatively long-lived insects. Here, we investigate the effect of post-emergence age on pheromone production in *Heliconius atthis* and *Heliconius charithonia* from two different tissue types: the wing androconia, responsible for male sex pheromones used in courtship behaviours, and the genital tip, responsible for the production of antiaphrodisiac pheromones that affect post-mating behaviour. These species have different mating strategies, with *H. atthis* being free-mating and *H. charithonia* being pupal-mating, which are hypothesised to affect the timing of their pheromone emissions . Gas chromatography-mass spectrometric analysis of androconial and genital extracts in males and females of both species from day 0 to 8 post-emergence demonstrates the following. Some ubiquitous non-pheromonal compounds and fatty acid precursors are already present at day 0. The complexity of the chemical blends produced by both tissue types in both sexes increases with age, often much earlier than day 8. No obvious difference in the time course of pheremone blend production was avident in these two species with contrasting mating strategies. While this analysis sheds light on some of the complexity in *Heliconius* pheromone production, much remains unknown about the effect of other factors such as the time of day or the presence of conspecifics.

Introduction

All living organisms utilize signals to communicate with and elicit responses from other individuals. When it comes to animals, the signals may convey different messages (e.g., readiness to mate, aggression, aggregation) and come in a number of different forms depending on which aspect of the recipient's perception they target (e.g., visual, acoustic, olfactory), and the recipient itself may be conspecific or heterospecific to the signaller (Laidre and Johnstone, 2013). Regardless of their function, however, something most signals have in common is that they are heavily context dependant, as the physiological state of the signaller, that of the receiver, and even the external environment may affect how information is conveyed and how it is perceived (Endler, 1992; Laidre and Johnstone, 2013). For this reason, the timing of the signals is important in order for information to be appropriately interpreted, and most species do not signal continuously but rather transmit under specific conditions and at specific moments of their life cycle (Endler, 1992).

Chemical signals are widespread among living organisms, ranging from bacteria to plants, to fungi and animals both vertebrate and invertebrate (Hardege *et al.*, 1998; Snell, 1998; Tillman *et al.*, 1999; Gosling and Roberts, 2001; Casselton, 2002; Stacey, 2003; Jacob *et al.*, 2004; Grammer, Fink and Neave, 2005; Sekimoto, 2005; Luporini, Alimenti and Vallesi, 2015; Dong *et al.*, 2020). Among them we find pheromones (Wyatt, 2009). Defined (Karlson and Lüscher, 1959) as blends of behaviourally active compounds that mediate communication between conspecific or often even heterospecific individuals, pheromones are among the most important signals in the animal kingdom, mediating several decisions and behaviours adopted by the organism during its life cycle (Yew and Chung, 2015). Pheromone composition is generally species-specific (Yew and Chung, 2015), or even race-specific in some known cases (Kochansky *et al.*, 1975; Glover, Tang and Roelofs, 1987; Bengtsson *et al.*, 2014). This specificity is not only derived from the identity of the blend components, but also their ratios, with receiving individuals being able to perceive minute differences in composition and titre (Klun *et al.*, 1973; Glover, Tang and Roelofs, 1987). On top of communicating the species identity of the signalling individual to all the receivers, pheromones may have a function in courtship, mate choice, aggregation, aggression and even social structure (Johansson and Jones, 2007; Wyatt, 2009; Yew and Chung, 2015).Thus, the context in which pheromones are emitted and the physical conditions of the signalling individual are extremely important. Insects have been shown to be able to discern a conspecific's age, mating status and sex based on their chemical signatures (Cuvillier-Hot *et al.*, 2001; Everaerts *et al.*, 2010; Nieberding *et al.*, 2012; Heuskin *et al.*, 2014; Polidori *et al.*, 2017), and the timing

and composition of pheromone emissions themselves is regulated by factors such as age and circadian rhythms, as well as feeding regime (Raina, Klun and Stadelbacher, 1986; Rafaeli and Soroker, 1989; Larsdotter-Mellström *et al.*, 2012; Nieberding *et al.*, 2012; Heuskin *et al.*, 2014; Ruther *et al.*, 2014; Darragh, Byers, *et al.*, 2019). This means the chemicals potentially carry large amounts of information that the conspecifics are able to decode and interpret (Johansson and Jones, 2007).

In Lepidoptera (moths and butterflies), pheromones are widely used in mating (Roelofs and Rooney, 2003; Yew and Chung, 2015). Female moths emit long distance sex pheromones that attract mate-seeking males (Zhang and Löfstedt, 2015; Allison and Carde, 2017), whereas in butterflies, males produce volatile male sex pheromones (MSP) during courtship (Myers, 1972; Nieberding *et al.*, 2008). Female butterflies perceive these signals, and thanks to the chemical specificity, use them to assess the male's species identity, which in turn influences their decision to mate. Another class of pheromones some species of butterflies can produce is antiaphrodisiacs. These are physically delivered from males to females through their genitals during mating, and their purpose is to signal the mated status of the female and reduce competition and harassment from other males (Andersson, Borg-Karlson and Wiklund, 2000, 2003; Schulz *et al.*, 2008; Estrada *et al.*, 2011). Previous studies have shown that female butterflies can utilize the males' pheromone composition to discern their age (Drickamer and Brown, 1998; Ting, Kelly and Snell, 2000; López, Aragón and Martín, 2003; Johansson and Jones, 2007; Coppée *et al.*, 2011; Karl, Heuskin and Fischer, 2013), and in many cases it has been shown that this information is taken into account during mate choice (Drickamer and Brown, 1998; López, Aragón and Martín, 2003; Coppée *et al.*, 2011; Nieberding *et al.*, 2012; Karl, Heuskin and Fischer, 2013).

Heliconius butterflies are regarded as model organisms in evolutionary biology, and research on their chemical signals has only begun garnering more interest in the past decade, with studies on both their MSPs, emitted during courtship from specialized scales on their wings known as androconia (Vanjari *et al.*, 2015; Darragh *et al.*, 2017, 2020; Mann *et al.*, 2017), and their antiaphrodisiacs (Schulz *et al.*, 2008; Estrada *et al.*, 2011; Darragh, Orteu, *et al.*, 2019; Byers *et al.*, 2020). *Heliconius* blends include several types of component, including plant derivatives and fatty acid (FA) derivatives, produced from mainly three fatty-acid precursors (palmitic, stearic and oleic acid) via a metabolic pathway that is conserved among Lepidoptera (Tillman *et al.*, 1999; Liénard *et al.*, 2014; Yew and Chung, 2015). Female *Heliconius* are known to participate actively in mate choice, which involves more than being able to tell conspecific and heterospecific males apart. While female *Heliconius* can use pheromone blends to recognize members of their own species (Darragh *et al.*, 2017; Southcott and Kronforst, 2018; González-Rojas *et al.*, 2020), they can also choose to reject or accept copulation from conspecific

males, a decision which is believed to be mediated by MSPs through as-yet poorly explored mechanisms.

Understanding the effect of endogenous (e.g. age, mating status, host plant) and exogenous (e.g. time of the day, temperature, presence/absence of conspecifics or competitors) factors on male *Heliconius* pheromone composition may prove crucial to uncovering how females are interpreting these signals. To further complicate the topic, *Heliconius* butterflies are known to engage in two different mating strategies. Many species are "free-mating" and commence copulation after male courtship. But in species belonging to the "pupal-mating" clades (Gilbert, 1976; Deinert, Longino and Gilbert, 1994; Beltrán *et al.*, 2007), males are attracted to female pupae and mate with females soon after or sometimes even before eclosion, with no apparent female choice involved (Beltrán *et al.*, 2007; Estrada *et al.*, 2010; Thurman *et al.*, 2018). Even though this behaviour is facultative, with pupal-mating species not losing the ability to court and mate as adults, (Thurman *et al.*, 2018), it may have measurable effects on the timing of pheromone production in these species, implying potentially less investment from males in pheromone production as no courtship is involved. Pupal-mating males mainly compete amongst themselves to land on female pupae and guard them, thus gaining the possibility to mate as soon as the female emerges (Deinert, Longino and Gilbert, 1994), and while the pupae themselves emit chemical cues to attract males (Estrada *et al.*, 2010), females cannot actively partake to mate choice and once copulation is complete they are very unlikely to re-mate (Walters *et al.*, 2012). The reduced post-mating competition (Estrada *et al.*, 2011) may mean that pupal-mating males would potentially invest less than free-mating males on antiaphrodisiacs, with their volatile genital contents showing no correlation with the onset of sexual maturity and fewer differences in composition between sexes. Conversely, freshly emerged female tissues might be expected to contain volatiles related to the pupal-mating behaviour, potentially the same described in (Estrada *et al.*, 2010). A similar trend might be expected in male androconia: as pupal-mating males would engage in courtship less frequently than free-mating males, the androconia might contain fewer unique and age-dependent compounds, not found in control tissues, compared to those of free-mating species.

As it has been demonstrated that age is one of the biological factors that most commonly affects pheromone production, knowing the effect of this parameter on *Heliconius* chemical emissions is important and has the potential of being useful for future experimental designs centred around courtship behaviours and physiological changes associated with sexual maturity and mating. Importantly, this information may provide information on how females discriminate between different conspecific males when it comes to mating. Furthermore, exploring the levels of different compounds may provide insight into how pheromone biosynthesis is carried out based on whether or

not the precursors are accumulated within the same tissues as the products as is the case for other Lepidopteran species (Liénard *et al.*, 2014).

Here, we investigate the effect of age on the volatile chemical contents of two male *Heliconius* pheromone-producing tissues: the androconia responsible for MSPs (Vanjari *et al.*, 2015; Darragh *et al.*, 2017; Mann *et al.*, 2017) and the genitals, which produce antiaphrodisiac pheromones (Gilbert, 1976; Schulz *et al.*, 2008; Klein and de Araújo, 2010; Estrada *et al.*, 2011), alongside female samples from analogous tissues. I undertook this investigation in the free-mating species *Heliconius atthis* and the pupal mating species *Heliconius charithonia* (Estrada *et al.*, 2010; Walters *et al.*, 2012) to answer the following questions: 1) Is age a good predictor of pheromone composition, and how does it affect the number and type of compounds produced? 2) How many and which compounds are age-dependent and does this vary with sex? 3) Do precursor compounds appear before the final blend components? 4) Is there any difference in the effect of age between the two different *Heliconius* mating strategies, given their different mating dynamics?

Materials and methods

Specimen and tissue extraction

All specimens used in this study were purchased from Stratford Butterfly Farm (Stratford, United Kingdom) as pupae, sourced from populations in Ecuador, Colombia and El Salvador. The chosen species were *Heliconius atthis* for the free-mating clade and *Heliconius charithonia* for the pupal-mating clade. On eclosion, all adults had their wings marked to allow individual identification. Adult butterflies were housed in a large cage at the University of York (UK) greenhouses at a temperature of at least 25°C, and fed on a mixture of water, honey and pollen.

The sampling protocol was the same for individuals from the two species and sexes. Based on courtship and mating activity, male *Heliconius* are known anecdotally to reach sexual maturity around ~7-8 days post-eclosion (Rosser & Mallet 2021, Personal Communication, 20 Dec.), so individuals were sacrificed at five different ages: 0, 2, 4, 6 and 8 days after eclosion. For each time point we sampled between 2 and 5 individuals of each sex, and from each individual we dissected three different tissue samples: 1) the wing androconia responsible for emitting the courtship pheromones (Darragh *et al.*, 2017; Rosser *et al.*, 2019) (also see Chapter 1) and equivalent regions from the females; 2) a control region of the wing; 3) the last abdominal segment: in males, this is where claspers are located, which carry the antiaphrodisiac producing glands, whereas females possess complementary structures to the claspers wherein these scent bouquets are deposited (Eltringham, 1925; Schulz *et al.*, 2008; Estrada *et al.*, 2011). The samples were dissected using sterile forceps and scissors and immediately suspended in 200 µL of a solution of dichloromethane and 2-acetoxytetradecane at a concentration

of 1 ng/µL to act as the internal standard (IS), in a 2 mL glass vial. The IS was used to calculate the amounts of other compounds (nmol). Two of the day 6 *H. charithonia* females had previously mated, but were included in the study as their genital contents showed no difference in composition to unmated females.

Gas Chromatography-Mass Spectrometry (GC-MS) settings and data processing

Tissue extracts from both sexes were analyzed on a GC-MS system (Department of Chemistry , University of York) which consisted of a 7890A GC-System (Agilent Technologies, Santa Clara, CA, USA) coupled with a Waters GCT Premier TOF Mass Analyzer (Waters Corporation, Milford, MA, USA) fitted with a Phenomenex ZB5-MSplus (30 m x 0.25 mm x 0.25 µm) column (Phenomenex, Macclesfield, UK). To improve mass accuracy, MS grade perfluorotributylamine (Heptacosa, Code: PC0568; Apollo Scientific Ltd., Stockport, UK) was utilized as a constant-flow calibrant. The ionization method was electron ionisation with an electron energy of 70 eV. The instrument settings were the following: inlet pressure 9.79 psi, He 20 mL min^{-1}, injection volume 1 µL. The GC was programmed with the following temperature ramp: starting at 50°C, increased by 5°C min^{-1} to a maximum temperature of 320°C, then held for 5 min for a programme total of 69 min. The samples were analyzed in batches alongside a C6-C40 alkane ladder prepared at Braunschweig TU (Braunschweig, Germany). This alkane ladder was used to calculate the retention index (RI) for each compound, necessary for identification and comparison with existing databases, particularly with the NIST Mass Spectra and Retention Indices databases (Linstrom and Mallard, 2021).

Data processing

The GC-MS data files were converted into .CDF format and processed via untargeted analysis using the Python package PyMassSpec (Davis-Foster, 2019, originally from O'Callaghan *et al.*, 2012). Compound identification was later carried out in the AMDIS software (Stein, 1999) with the aid of libraries compiled at Braunschweig TU (Braunschweig, Germany). All analyses were run in R using the base stats package. Genital extract data were analyzed separately from wing tissue (androconia and control) extract data. Thus, 6 datasets were generated for each species (12 in total): male androconia, male androconia control, male genitals, female "androconia", female "androconia" control and female genitals.

Data analysis

To test whether age is a good predictor of pheromone composition, the dimensionality of the data was first reduced using the metaMDS function in the R(vegan) package on each of the six tissue types per species, with k=2 and based on 1000 permutations of the data. A linear discriminant analysis (LDA) model was then run on the two resulting NMDS axes, NMDS1 and NMDS2 to test the extent to which

global pheromone composition was different between each age class in each of the six different tissue types per species (male androconia, male control, male genitals, female androconia, female control, female genitals).This model was trained using a partition of 60% of the dataset and its accuracy was calculated using the jackknife (leave-one-out) cross-validation approach, chosen due to the small sample size.

A linear model was used to test the effect of age on each individual compound's amount separately for the six tissue types per species. The resulting p-values were corrected using the Benjamini-Hochberg procedure for controlling the False Discovery Rate (Benjamini and Hochberg, 1995). A one-way ANOVA was used to test for changes in the number of detected compounds over time, using number of compounds as the dependent variable and individual age, in this case treated as a categorical variable, as the independent variable. This was carried out for each tissue type and for each species separately. The results of this test were further explored with Tukey's HSD test, in order to discern the specific time points at which different compounds were acquired.

To identify the signature compounds of each tissue type, I ran an Indicator Compound Analysis (ICA) using the R(indicspecies) package on each of the 12 tissue types separately to identify the indicator compounds, utilizing only day 8 samples. This analysis consists of a different application of the Indicator Species Analysis (ISA) described in (Dufrêne and Legendre, 1997; De Cáceres and Legendre, 2009; De Cáceres, Legendre and Moretti, 2010): where the analysis is normally utilized to uncover associations between species and sites, we use it here in an analogous manner to test the association between individual compounds and tissues, as done in (Heuskin *et al.*, 2014; Darragh *et al.*, 2020). The ICA, like the ISA, assigns two values to each compound, which can both take on values in the 0-1 range: specificity (A) and coverage (B). A indicates how unique the compound is to extracts of a specific tissue type, with 1 meaning it is only found within extracts from that type. B indicates how widespread the compound is within extracts of a specific tissue type, with 1 meaning it is found within all extracts of that type. An indicator value IV can be obtained from A and B, with its maximum value being 1. An IV=1 for a compound within a tissue type indicates that the compound is found in all samples of that type (B=1) and in no other types (A=1). Nonetheless an indicator compound does not necessarily require IV=1, and lower values are admissible.

Results

The number and identity of compounds detected, as well as their amounts, varied greatly between tissue and sex. On day 8 the genitals extracts of both species always contained more detectable compounds than the wing tissues (both androconia and control) (Fig. 4.1). The number of compounds detected in the genitals extract was comparable between the two species, with males on average

containing a greater variety of compounds than females. Meanwhile, *H. charithonia*'s wing tissue extracts, regardless of sex, contained a greater variety of detectable compounds than *H. atthis'* (Fig. 4.1B).

Global effect of age on chemical blend composition

In order to determine whether age had an effect on the composition of the tissue extracts, I first utilized NMDS on each species' six tissue types as a dimensionality reduction technique to condense the data into a limited number of axes. I carried out a LDA on the first two NMDS axes, which can be seen plotted in Fig. 4.2. In general, while chemical blend composition is a reasonable predictor of age, particularly in the androconia of both species, there are varying degrees of overlap in chemical content between different age classes for extracts of the same tissue, which can be observed in the NMDS plots, where it is illustrated by the overlap between points belonging to different age groups (Fig. 4.2). These overlaps are reflected by the relatively low levels of accuracy of the LDA models (Supp. Table 4.1). I found varying levels of predictability of the chemical composition based on age group, with the LDA models performing better on wing tissues than on genital tissues, and *H. atthis* female androconia in particular reaching a relatively high jackknifed accuracy of 0.85, equivalent to a 15% chance of incorrect classification of the sample into the wrong age group. The lowest accuracy (0.5) was observed in *H. atthis* male genitals and female genitals and in *H. charithonia* female genitals, with 50% chance of incorrect classification.

Compound amounts over time

Linear models that tested for the effect of age on the amount of each compound revealed that post-eclosion age strongly affects the amount of compounds. Age was found to have a significant effect on the amounts of individual compounds (Fig. 4.3, Supp. Table 4.1), which are referred to as age-dependent compounds. We found compounds whose amount increased over time in all examined tissues, and in both species. However, which compounds were affected by age varied greatly between the different tissues, the sexes and the species, with each tissue type having a number of age-dependent compounds whose amount was not age-dependent in other tissue types from the same species, and there were no apparent patterns in the type of compounds that displayed these patterns. Most correlations between compound amount and age were positive, with no compounds decreasing significantly over time aside from fatty acids in *H. atthis* female control tissues, which then become undetectable on day 2.

In *H. charithonia* wing tissues, male androconia and male control had more age-dependent compounds (12 and 13 respectively) than female androconia and female control (4 and 7 respectively). The amount of five compounds was found to be correlated to age in male androconia but not in the

other three tissues: the plant derivative hexahydrofarnesyl acetone, henicosane, docosene, tricosadiene and pentacosene. As their amount actively increases concurrently with the onset of sexual maturity in male androconia only, these five compounds could be assumed to be good candidates as putative pheromones. A single compound, octacosanal, was significantly affected by age in male control. No compounds were found to be significantly affected by age exclusively in female androconia or female control.

While male androconia and male control were found to have more age-affected compounds than the equivalent female tissues, male genitals and female genitals have comparable numbers of compounds whose amounts increase over time (22 in male genitals and 20 in female genitals), and most of these age-dependent compounds are in common between the sexes. However, certain compounds are only age dependent in one of the two sexes. In male genitals, these are (Z)-9-octadecen-1-ol, pentacosene, 11-methylheptacosane, benzyl salicylate and several complex esters, which could thus be suitable candidates for the role of antiaphrodisiacs, whereas in female genitals they are mainly long-chain alkanes (25C+), alongside a tocopherol isomer and an unknown derivative of syringic acid.

H. atthis male androconia extracts contained more age-dependent compounds (13 in total) than those of female androconia, male control and female control (three, two and three compounds respectively). 11 compounds were uniquely age dependent in male androconia, two in female androconia (ethyl (Z,Z,Z)-9,12,15-octadecatrienoate and stearic acid), a single one in male control (palmitic acid), and two in female control (octadecatrienoic acid and cholesterol). As with *H. charithonia*, age dependent compounds in male androconia could be good pheromone candidates.

In *H. atthis* genital extracts, the situation is very different from those of *H. charithonia*. In this species, female genitals have only 12 compounds that significantly increase over time and only two that do so in a sex-specific manner, the long chain alkane nonacosane and (Z)-13-docosenamide. Male genitals have a wide variety of age-dependent compounds (40 in total) that are either not detected in female genitals extracts or do not undergo an increase in amount, including the terpenes (E)-β-ocimene and alloocimene, several alkanes (both straight chain and branched), alkenes, esters and alcohols, as well as fatty acids. The presence of (E)-β-ocimene is especially noteworthy as this is a known antiaphrodisiac compound in some *Heliconius* species, wherein it also appears in combination with a complex blend of heavier compounds (Schulz *et al.*, 2008; Darragh, Orteu, *et al.*, 2019). Any other compounds that increase concurrently with the onset of sexual maturity could have a synergistic effect with (E)-β-ocimene, collectively creating an antiaphrodisiac effect.

Most of the age-correlated compounds across species, sexes and tissue types increase in amount over time, with only very few compounds showing a significant negative relationship between amount and

age. These are (Z,Z)-9,12-octadecadienoic acid (linoleic acid) in *H. atthis* male androconia, linoleic and stearic acid in *H. atthis* female androconia, palmitic, stearic and octadecatrienoic acid in *H. atthis* female control, cholesterol, phytyl acetate and linoleic acid in *H. charithonia* male androconia, linoleic and 9,12,15-octadecatrienoic acid, cholesterol and tocopherol in *H. charithonia* female androconia, cholesterol, phytyl acetate in *H. charithonia* male control. For a complete list of LM results including estimates, r^2 and p-values, see (Supp. Table).

Male *Heliconius* are thought to reach sexual maturity ~7-8 days after eclosion, therefore it is interesting that in *H. charithonia* male genitals and female genitals extracts, the majority of compounds with a significant correlation to individual age appear to reach maximum amount on day 6, to then decrease on day 8. In *H. atthis* male genitals and female genitals extracts, the amounts seem to continue increasing on day 8. Conversely, female *Heliconius* are believed to reach sexual maturity on day 0 after eclosion, so the fact that their tissues also contain volatiles that increase over time is unexpected, while intriguing.

Effect of age on number of compounds

In general, the number of different chemical species detected in the samples was found to increase from day 0 to day 8, however the pattern of this increase varied both with tissues and with species.

H. atthis *wing tissues:* In *H. atthis,* male control regions of the wing (male control) were the only tissue extracts that were found to not significantly increase the variety of individual compounds over time. In the male androconia, I found a significant effect of age on number of compounds (one-way ANOVA, $F(4,13)=18.07$, $p<0.0001$), and a significant difference in the number of compounds detected between day 0-2 (Tukey HSD, $p=0.004$), meaning most compounds would already be present in the tissues on day 2. All other days also showed significantly higher numbers of compounds than those detected in extracts of tissues on day 0, however for the purpose of this report we shall focus only on the comparisons of adjacent time points (2-0, 2-4, 4-6 and 6-8). See (Fig. 4.1) for a full representation of the results, showing changes in the number of compounds detected in the different tissues over the 5 time points. We found no significant difference in number of compounds between day 2-4, 4-6 and 6-8. Female androconia showed a similar trend to male androconia, with a significant effect of age on the number of compounds (one-way ANOVA, $F(4,15)=6.51$, $p=0.003$), and a significant difference between day 2-0 (Tukey HSD, $p=0.016$). For female control, I foud once again a significant effect of age (One-way ANOVA, $F(4,15)=5.99$, $p=0.004$) and a significant difference between day 2-0 ($p=0.014$) and 2-4 ($p=0.012$). However, for female control and female androconia, we found on average more compounds in day 0 extracts than those of day 2. The compounds detected on day 0 which then disappeared, becoming undetectable on day 2, were all fatty acids (palmitic acid, stearic acid and

linoleic acid) with the exception of a fatty acid ester, ethyl (Z,Z,Z)-9,12,15-octadecatrienoate. Female control then proceeds to gain a different small set of compounds on day 4, explaining the significant increase in the number of detected compounds. Nonetheless, the final number of detected compounds was still much higher in male androconia than in all three other wing tissue extracts (Fig. 4.1A).

H. atthis *genital tissues:* I found age to have a significant effect on the number of compounds in both male (male genitals) (one-way ANOVA, $F(4,13)=10.76$, $p=0.0004$) and female (female genitals) (one-way ANOVA, $F(4,15)=20.52$, $p<0.0001$) *H. atthis* genital extracts. In both cases there was no significant increase in number of compounds between adjacent time points but we detected significant increases when looking at larger time intervals: between day 0-6, 0-8, 2-6 and 2-8 in male genitals and between 0-4, 0-6, 0-8, 2-6 and 2-8 in female genitals (Fig. 4.1C).

H. charithonia *wing tissues: H. charithonia*'s wing tissues show a different pattern of compound accumulation from those of *H. atthis* (Fig. 4.1B). In *H. charithonia,* female androconia did not show any effect of time on the number of detected compounds (one-way ANOVA, $F(4,11)=2.15$, $p=0.142$). We found a significant effect of age on number of compounds in male androconia (one-way ANOVA, $F(4,16)=26.64$, $p<0.0001$), with a sharp increase between day 2-0 ($p=0.004$). There was a significant effect of age on male control as well (One-way ANOVA, $F(4,16)=10.61$, $p<0.0001$), though in this case we found no significant differences between adjacent time points, with the only ones showing significant differences being between day 6-0, 8-0, 6-2 and 8-2, hinting at a more gradual accumulation of particular components. Lastly, in female control, we see a significant effect of age (one-way ANOVA, $F(4,11)=18.01$, $p<0.0001$), and a significant increase between day 2-4 ($p=0.001$). Overall, all *H. charithonia* wing tissues displayed a greater variety of compounds than did *H. atthis* at all time points and in all wing tissues, with no significant decreases in the number of compounds at any time points.

H. charithonia *genital tissues:* For *H. charithonia* genitals, much like in *H. atthis,* we found a significant effect of age on the number of compounds in both male genitals (one-way ANOVA, $F(4,16)=9.05$, $p<0.0005$) and female genitals (one-way ANOVA, $F(4,11)=8.75$, $p=0.001$). In male genitals, there is a significant increase in number of compounds between day 0-2 ($p=0.04$) accompanied by several significant increases between non-adjacent time points, whereas in female genitals all significant increases were between non-adjacent time points (Fig. 4.1D).

Tissue specificity of compounds

To better discern potential pheromonal compounds and functionally relevant matrix components from structural components, I performed an Indicator Compound Analysis on data for extracts of every tissue separately, using only 8 days old individuals assuming that these carry the blend of

compounds characteristic of sexually mature adults. Indicator compounds of male androconia that are not found in other wing tissues of either sex may be good candidates as male sex pheromone components, and indicator compounds of male genitals that are not found in female genitals may be good antiaphrodisiac candidates. I successfully identified indicator compounds for several of the tissues featured in this study, reported in Table 4.1 alongside their specificity and coverage values. In *H. atthis* androconia (Table 4.1A), neither male or female wing control tissues had any indicator compounds. Female androconia only had 11-methylnonacosane, with an indicator value IV=0.866, specificity=1 and coverage=0.75, making it unique to that tissue but only detected in 75% of female androconia sample extracts. All of the other 17 detected indicator compounds were found in male androconia, with varying specificity and coverage values and indicator values ranging from 0.77 to 1, and of these 10 showed significant increases with age exclusively in this tissue. Among these compounds, the most prevalent class is alcohols (both saturated and unsaturated), alongside Z9-alkenes and aromatic compounds homovanillyl alcohol and syringaaldehyde. In *H. atthis* genitals (Table 4.1C), all 14 detected indicator compounds were from male genitals, all with complete coverage (B=1) and with A varying between 0.95 and, most often 1, and IV varying between 0.976 and 1 accordingly. All of these compounds were also significantly correlated with age in male genitals, except for hexyl (E,Z)-9,11-octadecadienoate. Once again, alcohols are prevalent in this tissue, accompanied by complex esters of fatty acids and by (E)-β-ocimene, a known antiaphrodisiac in *H. melpomene* (Schulz *et al.*, 2008).

Similar to *H. atthis*, in *H. charithonia* wing tissues (Table 4.1B) female androconia and female control had no indicator compounds, male control had 3 (11-methylnonacosane and two more unidentified nonacosane-derived branched alkanes) and male androconia had 11, albeit three of them (palmitic acid, stearic acid and cholesterol) had poor specificity, as they are detected in the extracts of the other three tissues as well. Several of the male androconia indicator compounds also had a correlation with age, and 6 of these showed this correlation exclusively in male androconia. The most prevalent type of compounds in *H. charithonia* male androconia is straight chain and methylated alkanes, accompanied by the plant derivatives hexahydrofarnesylacetone and phytyl acetate (the latter however does not seem to have a correlation with age). Lastly, in *H. charithonia* genitals (Table 4.1D), the indicator compounds are more evenly split between male genitals and female genitals. We find six indicator compounds in female genitals and eight in males, and most of them are also age dependent in the respective tissues. The *H. charithonia* male genitals indicator compounds included esters, a cyclic compound (2-(cyclohexylmethyl)butane-1,3-diol), and the plant derivatives neophytadiene and benzyl salicylate. Overall, there is little overlap between indicator compounds from different species and tissue groups, highlighting how each species' androconia and genitals have

their own distinct chemical signatures. Moreover, while many of these indicator compounds have shown significant correlations with age in one or more tissues, there are several more age-correlated compounds that do not show any tissue specificity.

Discussion

Age affects compound number and amount

Although the chemical composition of tissue extracts cannot be used to accurately predict age in these species due to overlap between different age classes (Fig. 4.2), the production of these blends appears to be strongly age dependent in *Heliconius* regardless of the mating strategy. Post-eclosion age affects both the amount of single compounds and the number of components detected in each tissue type. Production of compounds begins before sexual maturity, implying that the organism requires time to metabolize the pheromone components before reaching behaviourally active amounts. This is not unusual in butterflies, where male chemical blends are usually incomplete immediately after eclosion and reach the correct titre and composition as the adult ages and feeds (Nieberding *et al.*, 2008, 2012; Larsdotter-Mellström *et al.*, 2012; Darragh *et al.*, 2017). In most cases, day 8 compounds were already present on day 2 post-emergence, with very few first being detectable later than day 2. Regardless of the tissue or the sex, some compounds then begin increasing in amount until reaching their maximum at day 8 or day 6 in the case of *H. charithonia* male genitals, with the occasional introduction of new products in this span of time.

As this increase in compound amount also occurs in control wing tissues, and is thus not restricted to pheromone-emitting structures, one cannot assume that all compounds whose amount increases concurrently with the onset of sexual maturity are behaviourally active. It is likely that some of these age-dependent compounds found in negative controls, such as the many branched alkanes seen in *H. charithonia* (Fig. 4.34B, Table 4.1B) are cuticular components not involved in sexual signalling, as changes in cuticle composition with age have been demonstrated in other insects (Cuvillier-Hot *et al.*, 2001; Everaerts *et al.*, 2010; Heuskin *et al.*, 2014). Furthermore, fatty acid (FA) derivatives that make up *Heliconius* pheromone blends and cuticular hydrocarbons are produced via the same metabolic pathway (Liénard *et al.*, 2014; Chung and Carroll, 2015; Mann *et al.*, 2017), meaning its upregulation for the purpose of pheromone production as males approach sexual maturity may lead to higher levels of cuticular hydrocarbons as well.

Therefore, for a compound to be considered a good candidate as a male sex pheromone component, it would need not only to show a correlation between amount and age, but also to be found at significantly different levels in the genital or androconial extracts of males compared to females. As such, the male androconia and male genitals columns in Table 4.1 report what might potentially be

putative pheromone components or matrix components that enhance the effects of the behaviorally-active compounds (Schulz *et al.*, 2008). However, some of these compounds, like cholesterol and the tocopherol isomers, are rather heavy and so very unlikely to have a function as volatiles. This does not mean that age-dependent cuticular compounds have no function. For example, in *Bicyclus anynana,* cuticular hydrocarbons do indicate the sex and age of individuals, which are taken into account in mate choice (Heuskin *et al.*, 2014), and behavioural effects of cuticular hydrocarbons are not unusual (Howard and Blomquist, 1982; Chung and Carroll, 2015).

There is an interesting overlap in compounds that increase with age in female *B. anynana* and in female *H. charithonia* and *H. atthis*: all of them display higher relative amounts of alkanes in later days (6-8) post-eclosion compared to males (Heuskin *et al.*, 2014). It is possible that the age-dependent compounds detected in extracts of female tissues may have a role in short-range signalling, or even as contact pheromones (Würf *et al.*, 2020), a still unexplored topic as *Heliconius* female pheromones have received little attention. However, assuming the female age-dependent compounds are not just detected due to their aging cuticle, while intriguing, poses a problem. Female *Heliconius* reach sexual maturity soon after emergence, immediately in the case of *H. charithonia* due to pupal mating (Beltrán *et al.*, 2007; Estrada *et al.*, 2010), so why would their chemical signatures follow the same age-dependent pattern as that of males? If any of the female *Heliconius* compounds are involved in contact signalling, it may be deleterious to produce pheromones soon after emergence, potentially due to higher mortality in the early days and a greater investment being needed in survival rather than mating. Unfortunately, we do not know enough about female *Heliconius* chemical signatures outside the context of pupal mating to understand the reasons for this age effect.

Sample extracts do contain compounds on day 0 post-eclosion, probably derived from stored larval nutrients (Larsdotter-Mellström *et al.*, 2012). This is most noticeable in *H. charithonia* genitals, where on average 19(±7.3) and 17(±5.4) different compounds were detectesd respectively in male genitals and female genitals extracts from freshly eclosed individuals, compared to 11(±2.3) and 8(±5.2) in *H. atthis* day-0 male genitals and female genitals extracts respectively. These day-0 compounds are not sex specific in either species with the exception of a few alkanes in *H. atthis*. In fact, the two sexes' chemical signatures post-eclosion are extremely similar, becoming more differentiated as age increases. Among the day 0 compounds we find palmitic and stearic acids, alkanes, and in *H. charithonia,* very low amounts of compounds that are still detected in day-8 sample extracts. In both species these are accompanied by cholesterol and 1-3 isomers of tocopherol (including α-tocopherol or vitamin E), linoleic acid ((Z,Z)-9,12-octadecadienoic acid) and ethyl (Z,Z,Z)-9,12,15-octadecatrienoate. It could be assumed that these compounds may somehow be involved in pupal-mating functions, however none of them are *H. charithonia*-exclusive, and none of them are found

among pupal secretions of either *H. charithonia* sex (Estrada *et al.*, 2010). Although we cannot state whether they have a behavioural function or not due to the lack of studies of the matter, it seems more plausible that they may be related to energy storage and nutrition, as generally seen with sterols and polyunsaturated fatty acids in other insect species (Fraenkel and Blewett, 1945; Moadeli *et al.*, 2020).

Effect of age on genital contents

The genitals of both species show the richest variety of compounds (50 in total were detected in *H. atthis* when taking the data for the two sexes together, and 63 for *H. charithonia*) and the highest number of age-dependent compounds. In *H. charithonia* male genitals, most compounds reached the highest average relative amounts on day 6 post-eclosion, whereas in female genital extracts, this usually happened on day 8 (Fig. 4.3D). After the day 6 accumulation, male genitals appear depleted of volatiles. It may at first seem plausible that the accumulation may be a result of mating, as *H. charithonia* mated during the experiment and males may produce more compounds in proximity to mating and courtship behaviour, including antiaphrodisiacs (Schulz *et al.*, 2008; Estrada *et al.*, 2011). Even in species where male-produced antiaphrodisiacs have never been documented, it has been shown that mating does alter chemical signatures in insects, and that such changes are likely to affect mating attempts (Everaerts *et al.*, 2010; Polidori *et al.*, 2017). However, this phenomenon usually affects females, and it does not seem to be the case in our dataset: compounds that peak in 6-day-old *H. charithonia* individuals were detected on previous days too, showing a gradual accumulation over time.

This leaves us with the question of why this pattern is only observed in *H. charithonia* genital extracts and not those of *H. atthis,* where nearly all age-correlated genital compounds reach their maximum amount on day 8 regardless of sex. Unlike *H. charithonia, H. atthis* never mated during this experiment, potentially implying that the former may reach sexual maturity earlier, and that if we had collected *H. atthis* individuals at later time points, we might have been able to observe a peak and subsequent depletion of chemicals following mating in the genitals of the latter as well. The phenomenon of sex pheromone accumulation followed by a period of depletion in the first days of an insect's life cycle post-eclosion has been documented in sand flies and in moths (Raina, Klun and Stadelbacher, 1986; González *et al.*, 2017), however these species have a much shorter life cycle than the relatively long-lived *Heliconius,* that are able to survive for several months (Gilbert, 1972; Young and Montgomery, 2020). Nonetheless, it is possible that *Heliconius* pheromone production occurs in bursts and may be driven by other external abiotic and biotic factors (Baker and Cardé, 1979; Raina, Klun and Stadelbacher, 1986; Rafaeli and Soroker, 1989; Hagström *et al.*, 2012; González *et al.*, 2017).

Occurrence of fatty acid precursors

The fatty acid metabolic pathway that allows Lepidoptera to produce compound classes such alkanes, alkenes, alcohols, aldehydes and esters as part of its pheromones is based on three main fatty acid (FA) precursors: the saturated palmitic acid (hexadecanoic acid, C16:0) and its elongation product stearic acid (octadecanoic acid, C18:0) and the unsaturated oleic acid ((Z)-9-octadecenoic acid, C18:1), obtained via Δ9 desaturation of stearic acid (Liénard *et al.*, 2014; Mann *et al.*, 2017). Using two species that produce both saturated and unsaturated FA (fatty acid) derivatives, here I was able to test whether these three precursors accumulate in the pheromone secreting tissues to facilitate their transformation into more species-specific combinations of compounds. If that were the case, we would be able to observe a spike in FA precursor amount followed by a depletion as they are transformed. Palmitic and stearic acid are detected in day 0 extracts of all tissues tested, of both species. This may seem important due to their role as precursors to many saturated compounds in the butterflies' chemical bouquets, however their amounts tend to stay constant or increase over time, only showing a pattern of depletion in control wing tissues, not involved in pheromone signalling. In addition, oleic acid, the unsaturated equivalent of stearic acid and the precursor to many unsaturated compounds (Mann *et al.*, 2017), is entirely undetected in the dataset at all ages, whereas its unsaturated derivatives are detected, and increase over time. We cannot rule out the hypothesis that pheromones may be at least in part produced within the androconial and genital tissues however: the FA precursor, rather than accumulating all at once to then subsequently be depleted, may be constantly supplied to the pheromone-emitting tissues, rising and then reaching a plateau at the point where an equilibrium would be reached between the rate of accumulation and the rate of depletion, a pattern potentially more consistent with the observed results (see palmitic, stearic and (Z,Z)-9,12-octadecenoic acid in Fig. 4.3).

Thus, we have not found any conclusive evidence that the FA components of the chemical blends in *Heliconius* androconia and genitals are preceded by accumulation of FA precursors in these tissues, in contrast with findings on *Bicyclus anynana* where it has been demonstrated that biosynthesis does take place in the androconia (Liénard *et al.*, 2014). These compounds may instead be stored elsewhere in the organism, most likely in the fat body, haemolymph or oenocytes (Tillman *et al.*, 1999; Jurenka, 2004; Howard and Blomquist, 2005; Arrese and Soulages, 2010), to be later transported to other tissues for further transformation, as seen in moths (Schal, Sevala and Cardé, 1998; Jurenka *et al.*, 2003). In the case of genitals, this would be in contrast with what has been observed in *Bombyx mori* (Ohnishi *et al.*, 2011), where fatty acid pheromone precursors accumulate in pheromone glands on the tip of the abdomen. However, *Heliconius* butterflies, known for their unique pollen-feeding behaviour among butterflies and their long life cycle (Gilbert, 1972; Young and Montgomery, 2020),

may have a lipid allocation and/or transport mechanism unlike that of shorter-lived Lepidoptera that tend to subsist on nutrients acquired during the larval stage. It is also important to remember that while fatty acids are common pheromone precursors in *Heliconius* (Mann *et al.*, 2017), they are ubiquitous in animal tissues and function as energy storage in most organisms, so their presence in the organism soon after eclosion may be more indicative of that role rather than one in pheromone regulation.

Conclusions

In conclusion, I have demonstrated that aging has a strong effect on pheromone composition in *Heliconius.* We have found that while there is overlap between various age groups, which makes age a poor predictor of pheromone composition based on our LDA results, several compounds as well as the total number of individual chemical species show a striking correlation with age, with most compounds, whether cuticular or volatile, and regardless of sex, increasing in amount over time. The identity of these compounds and their number vary greatly across tissues and sex. We have also found that precursor compounds (palmitic, stearic and oleic acid) do not show a pattern of accumulation followed by depletion in pheromone-producing tissues, and are probably stored elsewhere in the organism. Even though palmitic and stearic acid are detected in all the analyzed tissue extracts, their amount either increases over time concurrently with that of all other compounds or stays relatively constant, showing no correlation with age. Lastly, we uncovered a few differences in the timing of pheromone production between *H. charithonia* and *H. atthis*, affecting the male *H. charithonia* genitals. We compared these two species as our aim was to observe whether pheromone production would follow a different pattern in the pupal-mating species where courting may be of reduced importance. While in every other tissue featured in this study, including the male *H. atthis* genitals, the compounds tend to reach maximum amount on day 8, in male *H. charithonia* genitals the peak is often reached on day 6. It is unknown why this difference in timing exists, but it is possible that *H. charithonia* reaches sexual maturity earlier than *H. atthis.* In order to determine whether this pattern is associated specifically with pupal mating species, a comparison involving more species from both the pupal-mating and the free-mating clades is required. Overall, aging revealed itself to be a complex topic in the context of *Heliconius* chemical ecology, and our work highlighted once more how little is known about pheromone perception, which chemical components are actually behaviourally meaningful and important in female choice, and what factors, external or internal, might be controlling their relative amounts. The prospect of further research on conditional pheromone production in these organisms, revealing progressively more and more of the information encoded in their chemical signals, is extremely fascinating.

Figures and Tables

A. *H. atthis* wing tissues	Female "androconia"		Male androconia		Female control		Male control				
	A	B	A	B	A	B	A	B	IV	p-value	Age effect
1-Octadecanol	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.015	AM
Methyloctadecan-1-ol	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.015	AM
(Z)-11-Icosenol	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.015	AM
Icosanol	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.015	AM
(Z)-13-Docosen-1-ol	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.015	AM
Unknown alcohol	0.089	0.75	**0.77**	**1.**	0.037	0.5	0.103	1.	0.879	0.005	AM-CM-CF
(Z)-9-Henicosene	0.	0.	**0.82**	**1.**	0.	0.	0.177	1.	0.907	0.015	AM
(Z)-9-Tricosene	0.	0.	**0.99**	**1.**	0.	0.	0.008	0.33	0.996	0.015	AM
11-Methylnonacosane	**1.**	**0.75**	0.	0.	0.	0.	0.	0.	0.866	0.02	n.s.
Ethyl (Z,Z,Z)-9,12,15-Octadecatrienoate	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.015	AF
Icosenylbutyrate	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.015	AM-CF
Methoxyhexadecane	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.015	AM
Palmitic acid	0.141	1.	**0.62**	**1.**	0.087	1.	0.157	1.	0.784	0.01	n.s.
(Z,Z)-9,12-Octadecadienoic acid	0.017	0.5	**0.98**	**1.**	0.	0.	0.	0.	0.991	0.01	n.s.
Stearic acid	0.023	1.	**0.81**	**1.**	0.024	0.5	0.138	1.	0.903	0.015	AF
Homovanillyl alcohol	0.	0.25	**1.**	**1.**	0.	0.	0.	0.	1.	0.02	AM
Syringaaldehyde	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.015	AM
Cholesterol	0.123	1.	**0.59**	**1.**	0.074	1.	0.21	1.	0.77	0.015	CF

B. *H. charithonia* wing tissues	Female "androconia"		Male androconia		Female control		Male control				
	A	B	A	B	A	B	A	B	IV	p-value	Age effect
Henicosane	0.	0.	**0.97**	**1.**	0.	0.	0.028	0.4	0.986	0.005	AM
Tricosane	0.029	0.67	**0.85**	**1.**	0.021	0.33	0.101	0.8	0.922	0.005	AM
Docosene	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.005	AM
Tricosadiene	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.005	AM
(Z)-9-Tricosene	0.011	1.	**0.94**	**1.**	0.009	1.	0.04	1.	0.969	0.005	all
Pentacosene	0.	0.	**0.98**	**1.**	0.	0.	0.018	0.2	0.991	0.005	AM
11-Methyltricosane	0.01	0.67	**0.81**	**1.**	0.022	0.33	0.161	0.8	0.899	0.015	AM-CM
11-Methylnonacosane	0.04	0.67	0.22	0.8	0.09	0.67	**0.645**	**1.**	0.803	0.03	AM-CM
Unknown branched alkane (RI 2952.6)	0.131	1.	0.22	1.	0.205	1.	**0.439**	**1.**	0.663	0.035	AM-CM
Unknown branched alkane (RI 2977.7)	0.	0.	0.	0.	0.365	1.	**0.635**	**1.**	0.797	0.005	CM
Palmitic acid	0.072	0.33	**0.56**	**1.**	0.169	0.67	0.195	1.	0.751	0.015	n.s.
9,12,15-Octadecatrienoic acid	0.	0.	**1.**	**1.**	0.	0.	0.	0.	1.	0.005	n.s.
Stearic acid	0.038	0.67	**0.49**	**1.**	0.174	0.67	0.298	0.8	0.7	0.01	n.s.
β-Tocopherol	0.	0.	**0.67**	**1.**	0.	0.	0.328	0.6	0.82	0.01	n.s.
Hexahydrofarnesylacetone	0.	0.	**1.**	**1.**	0.	0.	0.	0.2	1.	0.005	AM
Phytyl acetate	0.	0.	**0.82**	**0.8**	0.	0.	0.175	0.4	0.812	0.045	n.s.
Cholesterol	0.196	1.	**0.54**	**1.**	0.065	0.67	0.199	0.8	0.735	0.005	all

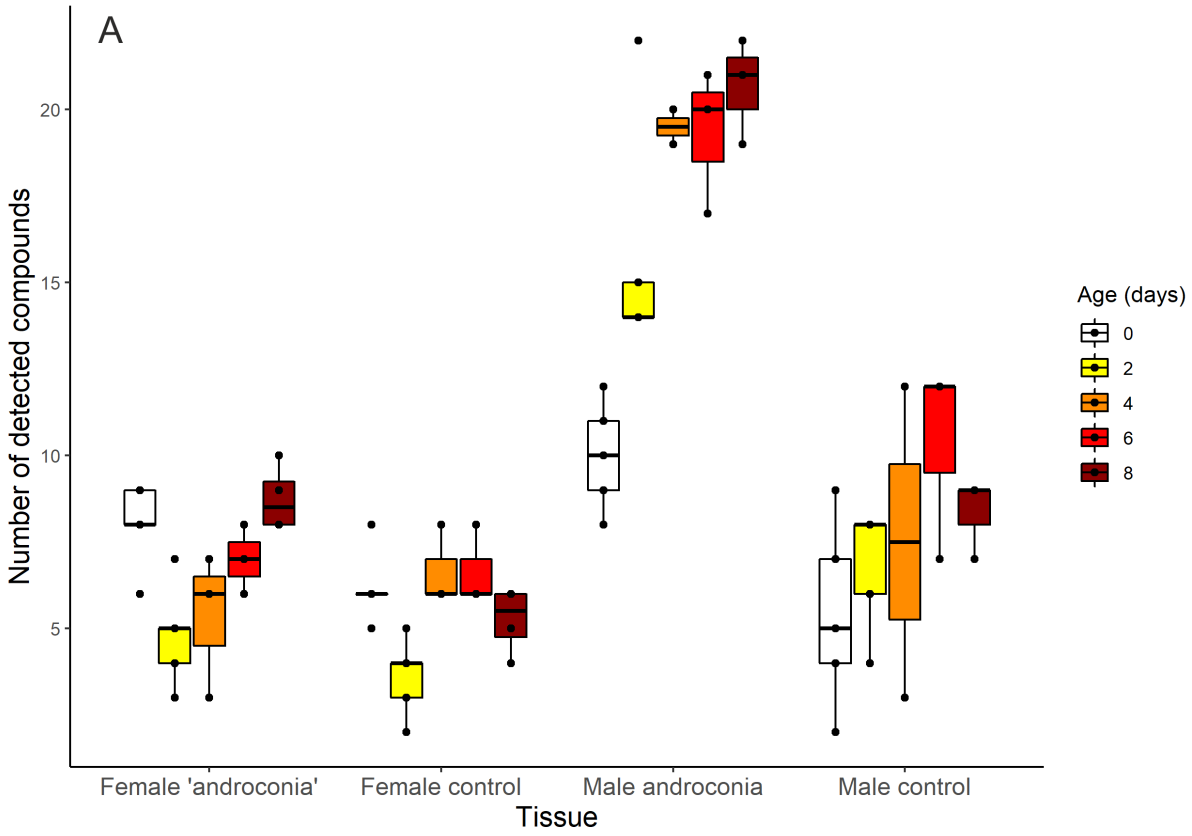
C. *H. atthis* genitals

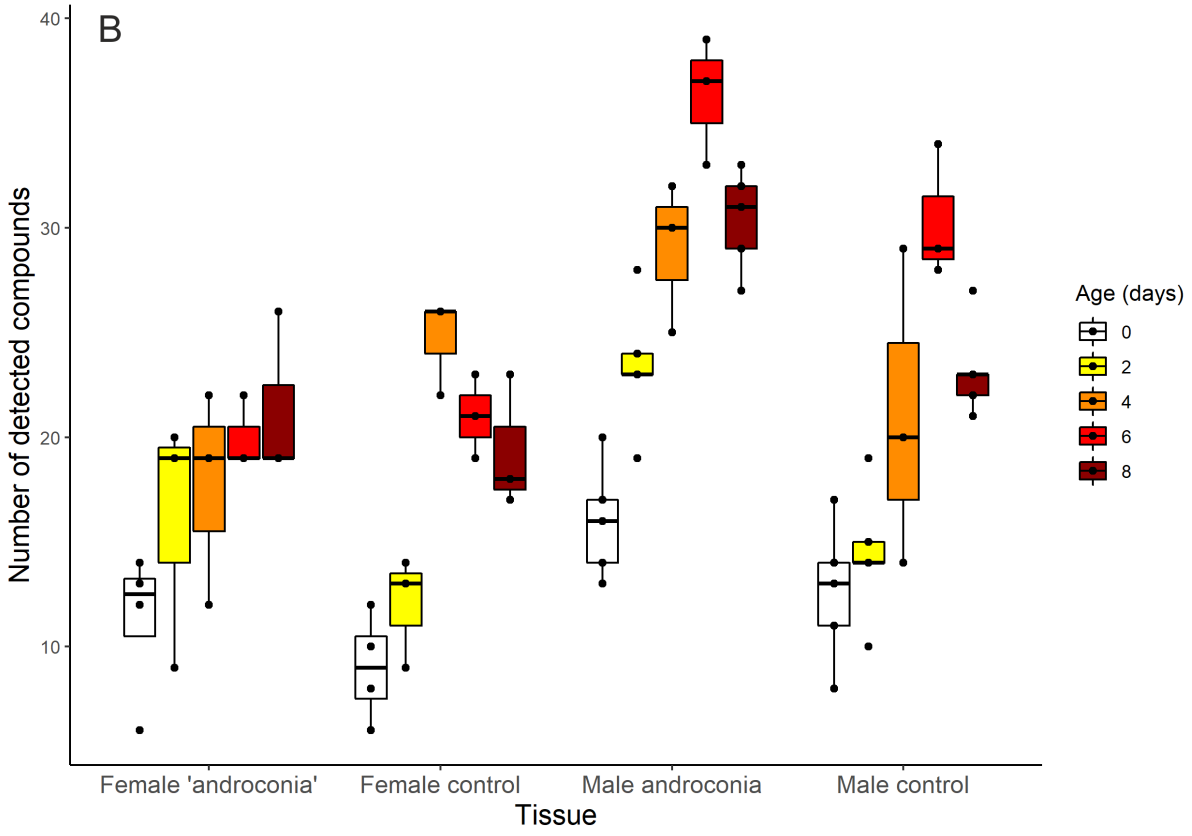
	Female genitals		Male genitals				
	A	B	A	B	IV	p-value	Age effect
1-Octadecanol	0.043	1.	**0.96**	**1.**	0.978	0.03	GM-GF
(Z)-11-Icosenol	0.	0.	**1.**	**1.**	1.	0.02	GM
(Z)-13-Docosenol	0.	0.	**1.**	**1.**	1.	0.02	GM
1-Tetracosanol	0.015	0.25	**0.99**	**1.**	0.993	0.03	GM
Unknown unsaturated alcohol (RI 2975.7)	0.	0.	**1.**	**1.**	1.	0.02	GM
(Z)-9-Henicosene	0.	0.	**1.**	**1.**	1.	0.02	GM
Docosene	0.	0.	**1.**	**1.**	1.	0.02	GM
Oleic acid long ester	0.044	0.25	**0.96**	**1.**	0.978	0.03	GM
(Z)-Hexenyl hexadecanoate	0.	0.	**1.**	**1.**	1.	0.02	GM
Hexyl (E,Z)-9,11-octadecadienoate	0.	0.	**1.**	**1.**	1.	0.02	n.s.
Hexyl octadecenoate	0.	0.	**1.**	**1.**	1.	0.02	GM
α-Tocopherol	0.048	0.75	**0.95**	**1.**	0.976	0.03	GM
3α,5-cyclo-5α-cholestan-6-β-ol	0.	0.	**1.**	**1.**	1.	0.02	GM
(E)-Ocimene	0.	0.	**1.**	**1.**	1.	0.02	GM

D. *H. charithonia* genitals

	Female genitals		Male genitals				
	A	B	A	B	IV	p-value	Age effect
2-(cyclohexylmethyl)butane-1,3-diol	0.	0.	**1.**	**1.**	1.	0.01	n.s.
Henicosane	**0.808**	**1.**	0.192	1.	0.899	0.04	n.s.
Heptacosane	**0.622**	**1.**	0.378	1.	0.789	0.04	GF
Nonacosane	**0.797**	**1.**	0.203	1.	0.893	0.025	GF
Hentriacontane	**0.854**	**1.**	0.146	0.8	0.924	0.025	GF
Pentacosene	0.119	0.67	**0.881**	**1.**	0.939	0.01	GM
11-Methylheptacosane	0.066	0.33	**0.934**	**1.**	0.966	0.01	GM
(Z)-3-Hexenyl hexadecanoate	0.	0.	**1.**	**1.**	1.	0.01	GM
(Z)-3-Hexenyl octadecenoate	0.	0.	**1.**	**1.**	1.	0.01	GM
Hexyl octadecenoate	0.	0.	**1.**	**1.**	1.	0.01	GM
α-Tocopherol	**0.973**	**1.**	0.027	0.4	0.986	0.03	n.s.
Neophytadiene	0.	0.	**1.**	**1.**	1.	0.01	n.s.
Benzyl salicylate	0.084	0.33	**0.916**	**1.**	0.957	0.01	GM
Cholesta-3,5-diene	**0.985**	**1.**	0.015	0.2	0.993	0.01	n.s.

Table 4.1. Indicator compounds for each tissue. A=specificity and B=coverage whereas IV=indicator value, taking both A and B into account. A indicates the uniqueness of each compound to that specific tissue (1=the compound is only found in that tissue), and B indicates the proportion of samples of that tissue the compound is found in (1=the compound is found in all samples). Bold values indicate the tissue for which the compound is a suitable indicator (high values of A and B). The p-value was obtained from 1000 permutations, and here we only report significant results. The age effect columns report whether age had a significant effect on that compound based on the linear models, and in what tissue; n.s. = non-significant relationship with age. Not all age-dependent compounds are indicator compounds and vice versa. AM=male androconia, CM=male control, GM=male genitals, AF=female "androconia", CF=female control, GF=female genitals.


A
Number of detected compounds
20
15
10
5
Female 'androconia'
Female control
Male androconia
Male control
Tissue
Age (days)
0
2
4
6
8


B
Number of detected compounds
40
30
20
10
Female 'androconia'
Female control
Male androconia
Male control
Tissue
Age (days)
0
2
4
6
8

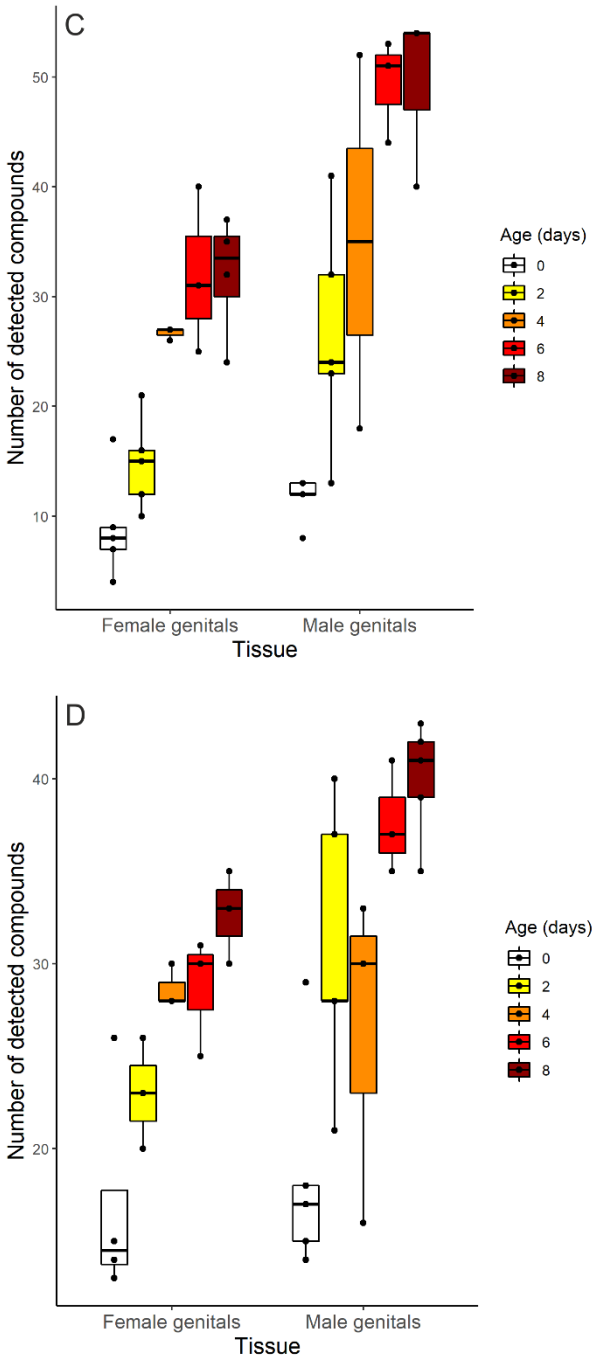
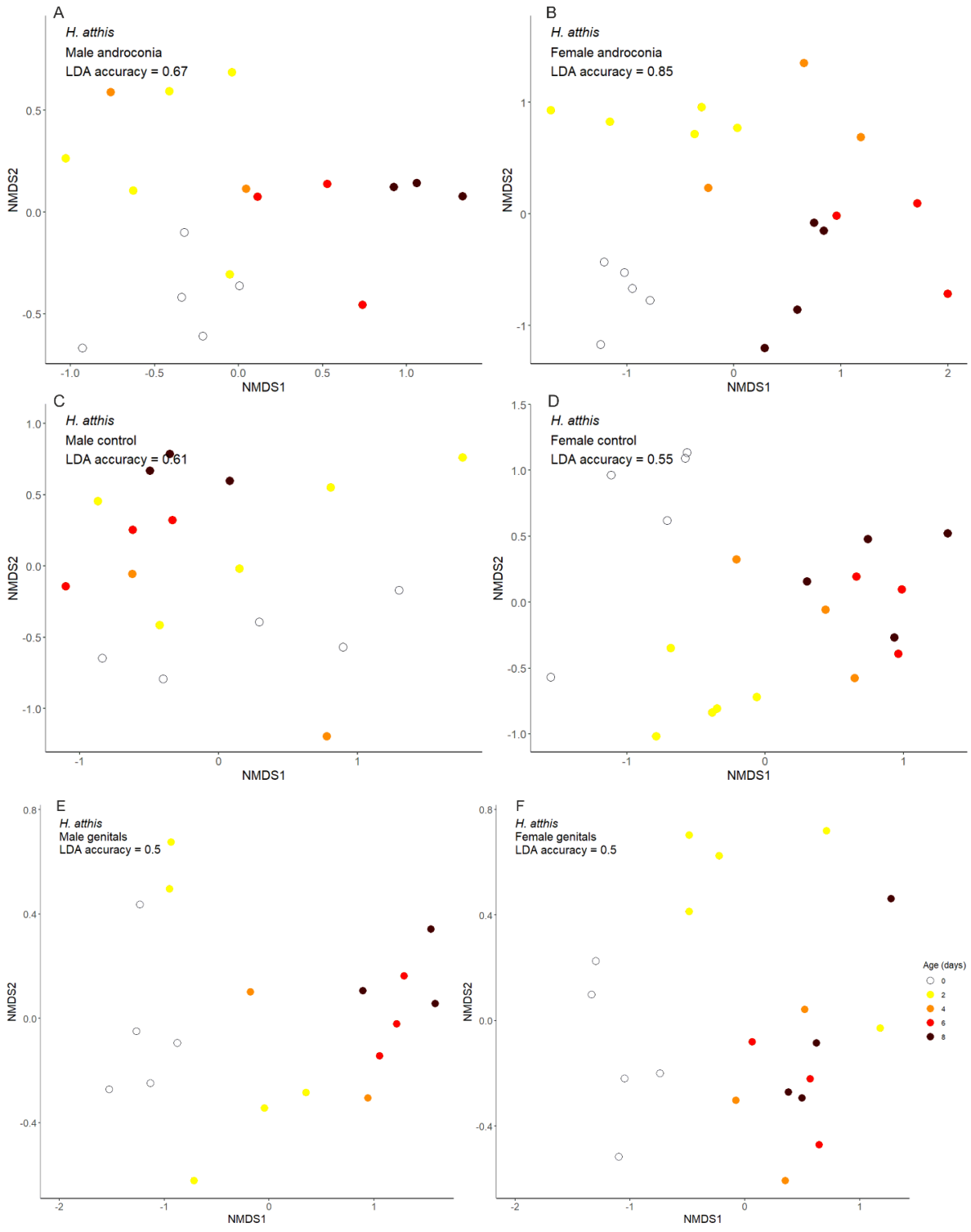


Figure 4.1. Number of individual compounds detected in extracts of *H. atthis* (A) and *H. charithonia* (B) wing tissues and in *H. atthis* (C) and *H. charithonia* (D) genitals at all measured time points.


A
H. atthis
Male androconia
LDA accuracy = 0.67
B
H. atthis
Female androconia
LDA accuracy = 0.85
C
H. atthis
Male control
LDA accuracy = 0.61
D
H. atthis
Female control
LDA accuracy = 0.55
E
H. atthis
Male genitals
LDA accuracy = 0.5
F
H. atthis
Female genitals
LDA accuracy = 0.5
NMDS1
NMDS2
Age (days)
0
2
4
6
8

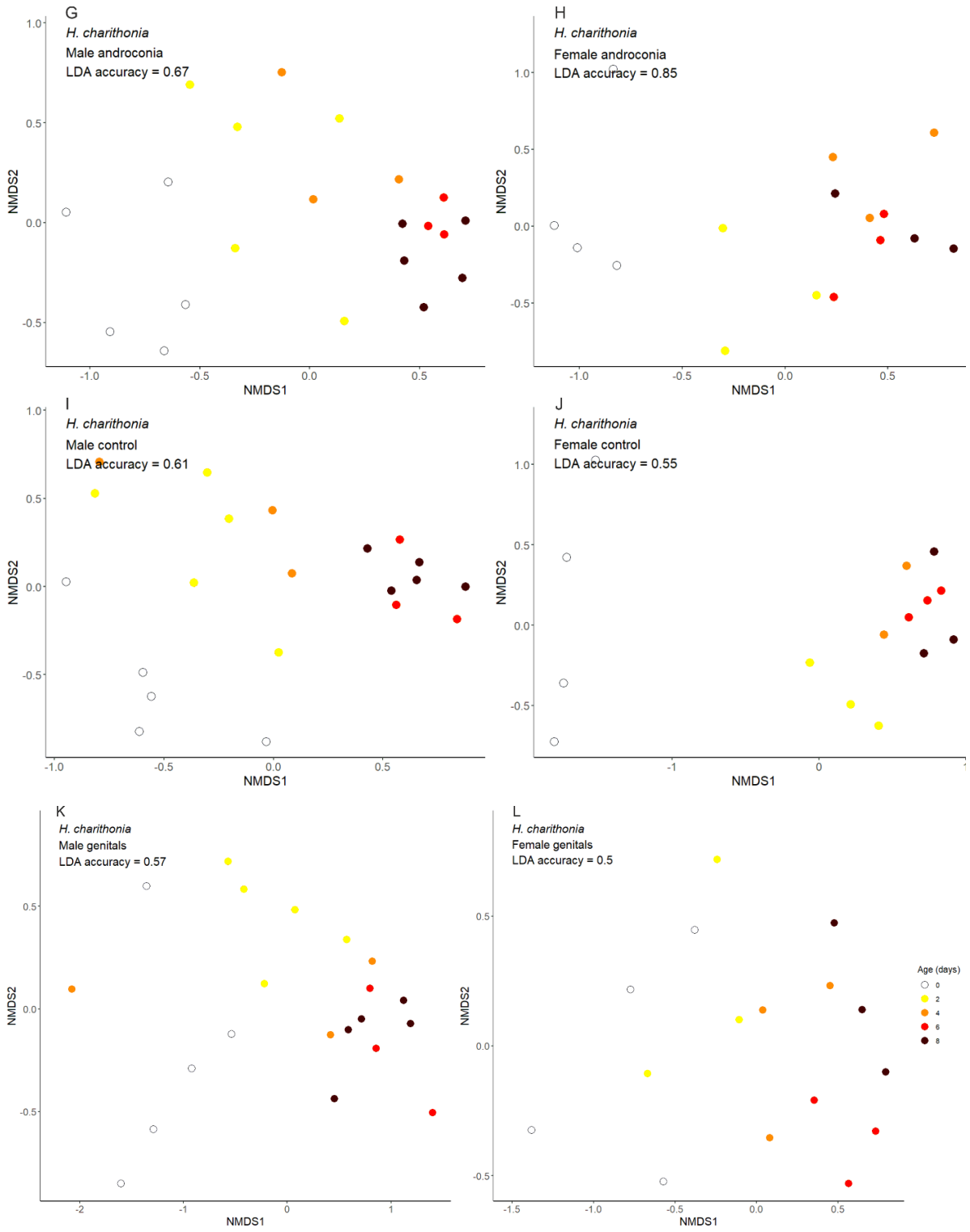
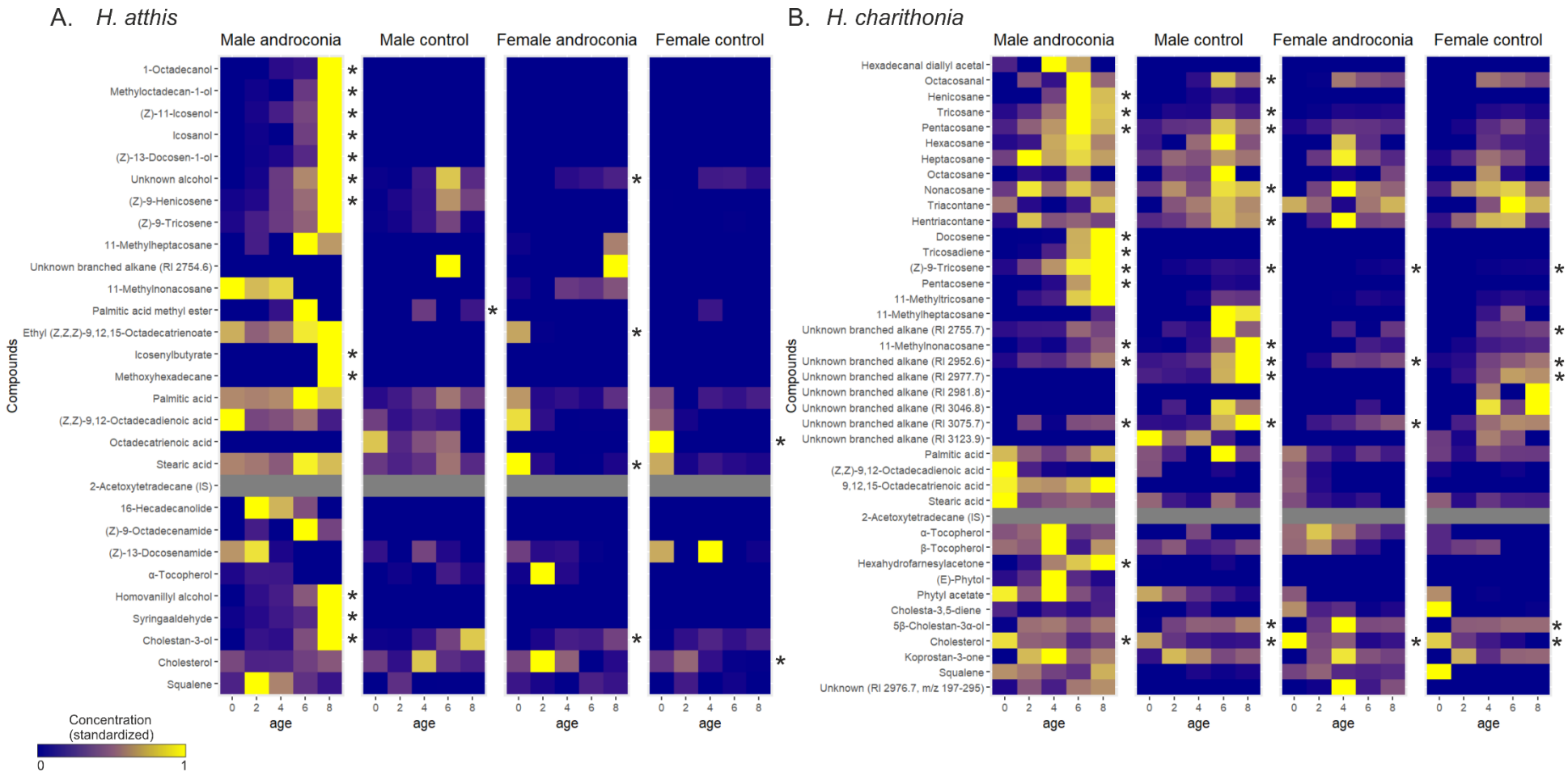


*Figure 4.2*A-L. NMDS plots based on k=2 and based on 1000 permutations of the data for the two species and all six tissue types (male androconia, male control, female "androconia", female control, male genitals, female genitals). Colors represent samples of the same age, with the 5 colours representing day 0, 2, 4, 6 and 8.


A. H. atthis
Male androconia
Male control
Female androconia
Female control
Compounds
1-Octadecanol
Methyloctadecan-1-ol
(Z)-11-Icosenol
Icosanol
(Z)-13-Docosen-1-ol
Unknown alcohol
(Z)-9-Henicosene
(Z)-9-Tricosene
11-Methylheptacosane
Unknown branched alkane (RI 2754.6)
11-Methylnonacosane
Palmitic acid methyl ester
Ethyl (Z,Z,Z)-9,12,15-Octadecatrienoate
Icosenylbutyrate
Methoxyhexadecane
Palmitic acid
(Z,Z)-9,12-Octadecadienoic acid
Octadecatrienoic acid
Stearic acid
2-Acetoxytetradecane (IS)
16-Hecadecanolide
(Z)-9-Octadecenamide
(Z)-13-Docosenamide
α-Tocopherol
Homovanillyl alcohol
Syringaaldehyde
Cholestan-3-ol
Cholesterol
Squalene
0 2 4 6 8
age
Concentration (standardized)
0
1
B. H. charithonia
Male androconia
Male control
Female androconia
Female control
Compounds
Hexadecanal diallyl acetal
Octacosanal
Henicosane
Tricosane
Pentacosane
Hexacosane
Heptacosane
Octacosane
Nonacosane
Triacontane
Hentriacontane
Docosene
Tricosadiene
(Z)-9-Tricosene
Pentacosene
11-Methyltricosane
11-Methylheptacosane
Unknown branched alkane (RI 2755.7)
11-Methylnonacosane
Unknown branched alkane (RI 2952.6)
Unknown branched alkane (RI 2977.7)
Unknown branched alkane (RI 2981.8)
Unknown branched alkane (RI 3046.8)
Unknown branched alkane (RI 3075.7)
Unknown branched alkane (RI 3123.9)
Palmitic acid
(Z,Z)-9,12-Octadecadienoic acid
9,12,15-Octadecatrienoic acid
Stearic acid
2-Acetoxytetradecane (IS)
α-Tocopherol
β-Tocopherol
Hexahydrofarnesylacetone
(E)-Phytol
Phytyl acetate
Cholesta-3,5-diene
5β-Cholestan-3α-ol
Cholesterol
Koprostan-3-one
Squalene
Unknown (RI 2976.7, m/z 197-295)
0 2 4 6 8
age

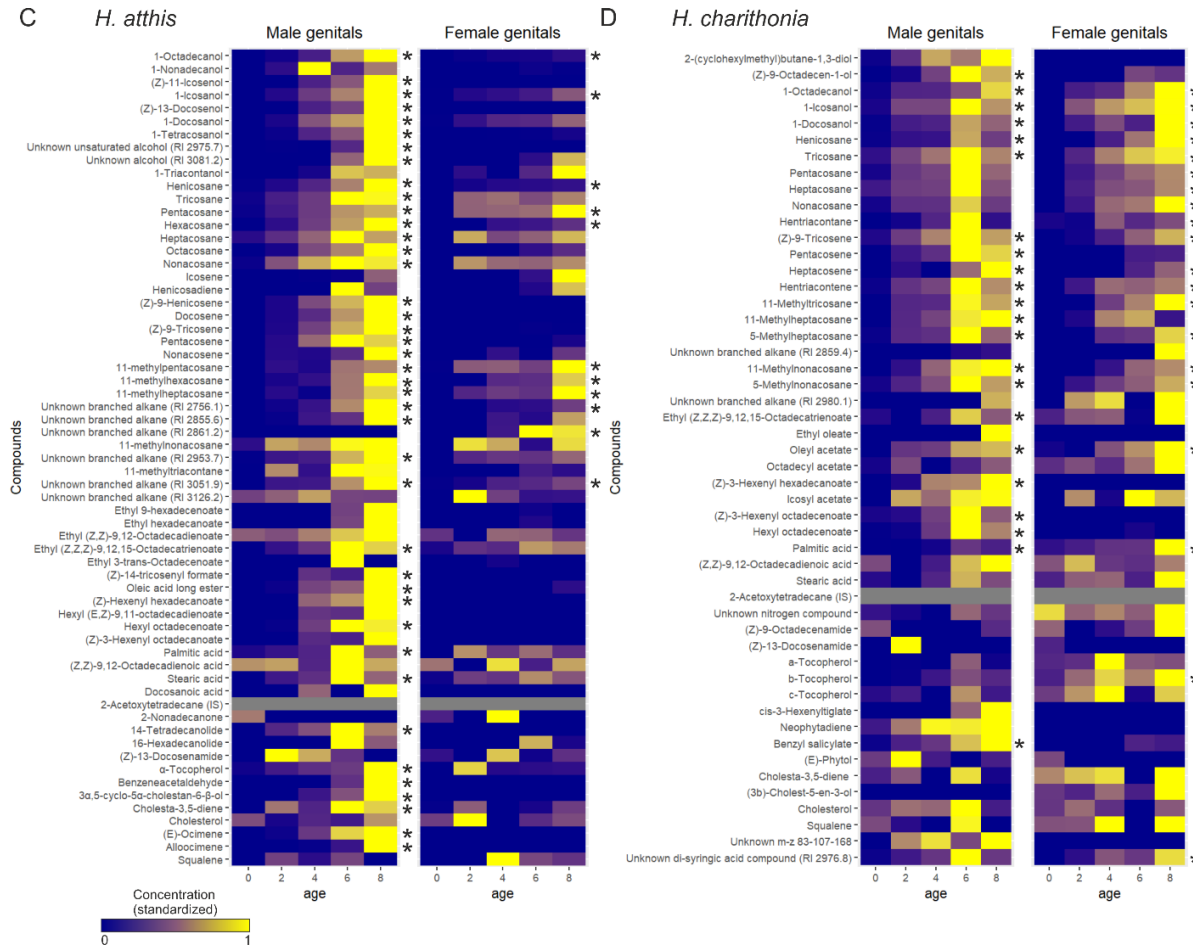


Figure 4.3. Changes in the relative amount with post-eclosion age for each compound detected in extracts of A) *H. atthis* wing tissues, B) *H. charithonia* wing tissues, C) *H. atthis* genitals and D) *H. charithonia* genitals. Compounds that show significant correlations with age are marked with an asterisk in each tissue type. 2-acetoxytetradecane (IS) is the internal standard we used to calculate each compound's amount, so its relative amount is constant (shown in grey). For each compound, the amount was standardized by dividing it by its maximum in that dataset. Within each class, compounds are ordered by retention index (top to bottom, low to high retention index).

Supplementary Material

H. atthis	Loadings	LD1	LD2		Proportion of trace	Accuracy (jackknifed)
Male androconia	NMDS1	3.167	0.507	LD1	0.86	0.67
	NMDS2	2.822	-3.046	LD2	0.14	
Male control	NMDS1	-0.903	-1.045	LD1	0.84	0.61
	NMDS2	1.995	-0.969	LD2	0.16	
Male genitals	NMDS1	-2.336	-0.323	LD1	0.96	0.50
	NMDS2	-3.149	2.764	LD2	0.04	
Female "androconia"	NMDS1	0.375	1.777	LD1	0.63	0.85
	NMDS2	2.618	-0.671	LD2	0.37	
Female control	NMDS1	2.536	0.303	LD1	0.90	0.55
	NMDS2	-0.398	1.705	LD2	0.10	
Female genitals	NMDS1	2.358	-0.775	LD1	0.68	0.50
	NMDS2	-2.241	-3.016	LD2	0.32	
H. charithonia						
Male androconia	NMDS1	5.215	0.187	LD1	0.94	0.63
	NMDS2	-1.258	-3.539	LD2	0.06	
Male control	NMDS1	-7.441	-0.948	LD1	0.91	0.67
	NMDS2	-3.004	2.359	LD2	0.01	
Male genitals	NMDS1	2.630	-0.347	LD1	0.94	0.57
	NMDS2	-0.676	-2.258	LD2	0.06	
Female "androconia"	NMDS1	4.283	0.330	LD1	0.94	0.60
	NMDS2	-0.339	2.324	LD2	0.06	
Female control	NMDS1	10.264	-0.349	LD1	1.00	0.69
	NMDS2	-0.599	2.454	LD2	0.00	
Female genitals	NMDS1	2.705	0.154	LD1	0.88	0.50
	NMDS2	-0.613	-2.835	LD2	0.12	

Supp. Table 4.1. LDA statistics on NMDS axes for both species and all tissue types. Loadings indicate each variable's contribution to the two LD functions, and "proportion of trace" refers to the proportion of data explained by each axis. The jackknifed accuracy of the LDA model is used to assess the fit of the model to the data and was obtained via leave-one-out (LOO) cross validation, ideal for smaller datasets.

A

Class	*H. atthis* Compound	Male androconia Estimate	R^2	FDR p-value	Male control Estimate	R^2	FDR p-value	Female androconia Estimate	R^2	FDR p-value	Female control Estimate	R^2	FDR p-value
Alcohol	1-Octadecanol	0.1443	0.5362	0.0089	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Methyloctadecan-1-ol	0.151	0.53	0.0089	NA	NA	NA	NA	NA	NA	NA	NA	NA
	(Z)-11-Icosenol	0.144	0.6153	0.0042	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Icosanol	0.1178	0.4621	0.0168	NA	NA	NA	NA	NA	NA	NA	NA	NA
	(Z)-13-Docosen-1-ol	0.1201	0.4781	0.0156	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Unknown alcohol	0.1323	0.5903	0.0042	0.1084	0.2853	0.7399	0.091	0.3089	0.0577	0.0079	0.3317	0.0457
Alkene	(Z)-9-Henicosene	0.1393	0.6017	0.0042	0.0808	0.2446	0.2693	NA	NA	NA	0.2301	0.0789	0.5932
	(Z)-9-Tricosene	0.1341	0.6299	0.0042	0.0405	0.0663	0.1454	NA	NA	NA	0.4291	0.0351	0.8159
Branched alkane	11-Methylheptacosane	0.0739	0.1887	0.252	NA	NA	NA	0.0502	0.1147	0.4177	NA	NA	NA
	Unknown branched alkane (RI 2754.6)	NA	NA	NA	0.0351	0.0495	0.7762	0.0509	0.1179	0.4111	NA	NA	NA
	11-Methylnonacosane	-0.0778	0.2039	0.2173	NA	NA	NA	0.08	0.2086	0.1607	NA	NA	NA
Ester	Palmitic acid methyl ester	0.0366	0.0543	0.7429	0.0405	0.0548	0.0113	NA	NA	NA	0.896	0.001	1.
	Ethyl (Z,Z,Z)-9,12,15-Octadecatrienoate	0.0744	0.1612	0.3269	NA	NA	NA	-0.1183	0.3879	0.0229	NA	NA	NA
	Icosenylbutyrate	0.1351	0.4496	0.018	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Methoxyhexadecane	0.1277	0.4347	0.0211	NA	NA	NA	NA	NA	NA	NA	NA	NA
Fatty acid	Palmitic acid	0.0639	0.1207	0.4463	0.1194	0.3568	0.3281	-0.0739	0.2092	0.1607	0.1703	0.1018	0.4703
	(Z,Z)-9,12-Octadecadienoic acid	-0.1045	0.2749	0.1095	-0.0704	0.1625	0.6624	-0.0918	0.3149	0.0555	0.0152	0.2855	0.0768
	Octadecatrienoic acid	NA	NA	NA	-0.0476	0.0809	0.7429	NA	NA	NA	0.0063	0.3468	0.0384
	Stearic acid	0.0495	0.0832	0.6061	0.025	0.0197	0.6528	-0.1166	0.4193	0.0168	0.0174	0.2757	0.0843
Macrolide	16-Hecadecanolide	-0.0099	0.0036	1.	NA	NA	NA	NA	NA	NA	NA	NA	NA
Nitrogen	(Z)-9-Octadecenamide	0.0492	0.0938	0.5706	NA	NA	NA	NA	NA	NA	NA	NA	NA
	(Z)-13-Docosenamide	-0.0468	0.0655	0.6946	-0.0095	0.0029	0.1483	-0.044	0.0871	0.5568	0.416	0.0371	0.8159
Phenol/ aromatic	α-Tocopherol	-0.0331	0.0397	0.8159	0.046	0.0854	0.6037	-0.0346	0.0357	0.8159	NA	NA	NA
	Homovanillyl alcohol	0.1638	0.7017	0.0017	NA	NA	NA	0.0509	0.1179	0.4111	NA	NA	NA
	Syringaaldehyde	0.1434	0.5159	0.0101	NA	NA	NA	NA	NA	NA	NA	NA	NA
Sterol	Cholestan-3-ol	0.1406	0.4619	0.0168	0.1727	0.6383	0.6446	0.1656	0.5417	0.0042	0.0197	0.2667	0.0916
	Cholesterol	0.035	0.0407	0.8159	-0.0022	0.0002	0.7429	-0.0681	0.1205	0.4111	0.0061	0.3485	0.0384
Terpene	Squalene	-0.022	0.0193	1.	-0.0236	0.0207	0.0984	0.0226	0.0177	1.	0.5322	0.022	0.9956

Acetal	Hexadecanal diallyl acetal	0.0164	0.0076	1.	NA	NA	NA	NA	NA	NA		NA	NA	NA	
Aldehyde	Octacosanal	0.0742	0.1968	0.1313	0.1471	0.4819	0.0045	0.1344	0.2916		0.1056	0.1324	0.351		0.0639
Alkane	Henicosane	0.14	0.5764	0.0012	0.0585	0.1383	0.2153	NA	NA	NA		NA	NA	NA	
	Tricosane	0.1775	0.6816	0.0002	0.1124	0.3445	0.0275	0.1087	0.207		0.1776	0.1263	0.2117		0.1767
	Pentacosane	0.1322	0.3955	0.0145	0.1152	0.3989	0.0142	0.1136	0.3612		0.0597	0.1001	0.2911		0.1056
	Hexacosane	0.1005	0.2339	0.0957	0.0785	0.1685	0.1639	0.0112	0.0046		1.	0.069	0.1093		0.3792
	Heptacosane	0.0213	0.0205	0.7939	0.0815	0.196	0.1313	0.0349	0.0369		0.7186	0.0497	0.0533		0.6172
	Octacosane	0.0302	0.0353	0.6441	0.0224	0.0177	0.83	0.0632	0.1443		0.2971	0.0077	0.002		1.
	Nonacosane	0.0134	0.0093	0.9804	0.106	0.3731	0.0187	0.0682	0.1664		0.2525	0.0774	0.1492		0.2899
	Triacontane	0.0368	0.0382	0.6211	0.0769	0.1153	0.278	-0.0005	0.		1.	0.1301	0.3309		0.0752
	Hentriacontane	0.0023	0.0003	1.	0.1381	0.3896	0.0153	0.0697	0.1117		0.3765	0.0662	0.0886		0.4548
Alkene	Docosene	0.1755	0.6891	0.0002	NA	NA	NA	NA	NA	NA		NA	NA	NA	
	Tricosadiene	0.1564	0.5381	0.0021	NA	NA	NA	NA	NA	NA		NA	NA	NA	
	(Z)-9-Tricosene	0.1913	0.6732	0.0002	0.1368	0.5141	0.0029	0.1309	0.5259	0.0103		0.2008	0.6236		0.0029
	Pentacosene	0.1057	0.4766	0.0047	0.0642	0.1581	0.1773	NA	NA	NA		NA	NA	NA	
Branched alkane	11-Methyltricosane	0.1046	0.4484	0.0076	0.0914	0.3133	0.0392	0.0659	0.1476		0.291	0.1155	0.2429		0.1397
	11-Methylheptacosane	0.0438	0.0964	0.3331	0.0712	0.1949	0.1313	NA	NA	NA		0.0342	0.042		0.6872
	Unknown branched alkane (RI 2755.7)	0.1019	0.2685	0.0645	0.0813	0.1903	0.1348	0.1114	0.2551		0.1314	0.1353	0.3973		0.0394
	11-Methylnonacosane	0.0945	0.333	0.0317	0.1201	0.5129	0.0029	0.0852	0.1885		0.2093	0.1078	0.2213		0.1648
	Unknown branched alkane (RI 2952.6)	0.142	0.5363	0.0021	0.1739	0.7045	0.0002	0.1498	0.527	0.0103		0.1265	0.4811		0.0171
	Unknown branched alkane (RI 2977.7)	NA	NA	NA	0.1944	0.6705	0.0002	NA	NA	NA		0.1899	0.732		0.0005
	Unknown branched alkane (RI 2981.8)	NA	NA	NA	NA	NA	NA	NA	NA	NA		0.0648	0.1283		0.3337
	Unknown branched alkane (RI 3046.8)	NA	NA	NA	0.0589	0.1327	0.229	0.0333	0.04		0.6974	0.1076	0.2664		0.1299
	Unknown branched alkane (RI 3075.7)	0.0506	0.0935	0.3346	0.1473	0.6203	0.0005	0.1713	0.5447	0.0088		0.125	0.4681		0.0191
	Unknown branched alkane (RI 3123.9)	NA	NA	NA	-0.0527	0.1164	0.2776	NA	NA	NA		-0.0101	0.0024		1.
Fatty acid	Palmitic acid	0.0057	0.0009	1.	0.0299	0.0248	0.7406	-0.0996	0.2076		0.1776	0.0122	0.0038		1.
	(Z,Z)-9,12-Octadecadienoic acid	-0.0863	0.2764	0.0605	-0.063	0.1595	0.1767	-0.087	0.2592		0.1313	-0.0415	0.0552		0.6097
	9,12,15-Octadecatrienoic acid	0.0147	0.0072	1.	NA	NA	NA	-0.0948	0.2928		0.1056	NA	NA	NA	
	Stearic acid	-0.0498	0.148	0.1944	-0.0171	0.0091	0.9804	-0.1181	0.3239		0.0796	-0.0595	0.1406		0.3011
Phenol	α-Tocopherol	-0.0116	0.0033	1.	0.002	0.0002	1.	-0.0447	0.0612		0.5867	-0.0556	0.1111		0.3765
	β-Tocopherol	0.0101	0.0034	1.	0.002	0.0001	1.	-0.1051	0.2737		0.1227	-0.057	0.0757		0.515
Plant	Hexahydrofarnesylacetone	0.183	0.6731	0.0002	0.0451	0.046	0.5846	NA	NA	NA		NA	NA	NA	
	(E)-Phytol	-0.001	0.	1.	NA	NA	NA	NA	NA	NA		NA	NA	NA	
	Phytyl acetate	-0.0824	0.2132	0.117	-0.0937	0.2601	0.0709	-0.0901	0.2477		0.1373	-0.0589	0.1257		0.3346
Sterol	Cholesta-3,5-diene	0.0011	0.	1.	0.002	0.0002	1.	-0.111	0.2449		0.1392	-0.0917	0.2255		0.1639
	5β-Cholestan-3α-ol	0.0719	0.1692	0.1639	0.1448	0.5583	0.0016	0.0385	0.06		0.5883	0.1435	0.4234		0.0317
	Cholesterol	-0.1042	0.3217	0.0355	-0.1191	0.4313	0.0093	-0.1031	0.3983	0.0394		-0.151	0.6145		0.0033
	Koprostan-3-one	0.0482	0.0754	0.406	0.0116	0.0066	1.	0.0426	0.0663		0.5652	0.0626	0.0881		0.4548
Terpene	Squalene	0.0043	0.0004	1.	0.0513	0.1065	0.2974	-0.0673	0.0912		0.4494	-0.0556	0.1111		0.3765
Unknown	Unknown (RI 2976.7, m/z 197-295)	0.0915	0.2012	0.1299	NA	NA	NA	0.045	0.0581		0.5957	NA	NA	NA	

Supp. Table 4.2. Linear regression results for all compounds in the wing tissues of A) *H. atthis* and B) *H. charithonia*. The p-values shown here were corrected according to the Benjamini-Hochberg procedure for controlling for the False Discovery Rate (FDR), to account for multiple testing (Benjamini and Hochberg, 1995). The estimate columns indicate the intercept, and R^2 the proportion of variance explained by the linear model. Highlighted cells indicate significant corrected p-values, and compounds that never show correlations with age are shown in grey.

A

Class	*H. atthis* Compound	Male genitals Estimate	R^2	FDR p-value	Female genitals Estimate	R^2	FDR p-value
Alcohol	1-Octadecanol	0.1987	0.7457	0.0001	0.1003	0.4139	0.01
	1-Nonadecanol	0.0585	0.1217	0.2397	0.0428	0.0789	0.3373
	(Z)-11-Icosenol	0.1773	0.6671	0.0003	NA	NA	NA
	1-Icosanol	0.1854	0.8084	0.	0.0781	0.2806	0.0387
	(Z)-13-Docosenol	0.1192	0.4669	0.0086	NA	NA	NA
	1-Docosanol	0.172	0.7157	0.0002	0.0714	0.2333	0.0685
	1-Tetracosanol	0.1105	0.462	0.0089	0.054	0.133	0.194
	Unknown unsaturated alcohol (RI 2975.7)	0.1004	0.3767	0.0203	NA	NA	NA
	Unknown alcohol (RI 3081.2)	0.1142	0.359	0.023	0.0523	0.1249	0.2093
	1-Triacontanol	0.0908	0.22	0.1008	0.0661	0.2015	0.099
Alkane	Henicosane	0.1441	0.5461	0.0028	0.1182	0.3545	0.0181
	Tricosane	0.1927	0.6167	0.0009	0.0788	0.1457	0.1794
	Pentacosane	0.209	0.8157	0.	0.1043	0.4059	0.0109
	Hexacosane	0.1758	0.6811	0.0003	0.0812	0.3083	0.0284
	Heptacosane	0.1761	0.5851	0.0016	0.0679	0.1279	0.2042
	Octacosane	0.1734	0.6926	0.0003	0.0606	0.1669	0.1386
	Nonacosane	0.1803	0.6777	0.0003	0.0607	0.1177	0.2157
Alkene	Icosene	0.075	0.2228	0.0991	0.0666	0.1943	0.1035
	Henicosadiene	0.0506	0.0997	0.3027	0.0798	0.2349	0.0683
	(Z)-9-Henicosene	0.1827	0.7947	0.	NA	NA	NA
	Docosene	0.1589	0.6286	0.0008	NA	NA	NA
	(Z)-9-Tricosene	0.1945	0.7463	0.0001	0.0388	0.0686	0.3831
	Pentacosene	0.1187	0.3736	0.0203	0.0509	0.1179	0.2157
	Nonacosene	0.0988	0.373	0.0203	0.0519	0.1221	0.2117
Branched alkane	11-methylpentacosane	0.1481	0.5719	0.002	0.083	0.3004	0.0307
	11-methylhexacosane	0.1179	0.3934	0.018	0.1224	0.5109	0.0025
	11-methylheptacosane	0.1271	0.5285	0.0036	0.101	0.4316	0.0083
	Unknown branched alkane (RI 2756.1)	0.183	0.6219	0.0008	0.0968	0.3864	0.0129
	Unknown branched alkane (RI 2855.6)	0.1325	0.4008	0.0168	0.0547	0.1377	0.1902
	Unknown branched alkane (RI 2861.2)	NA	NA	NA	0.1391	0.5107	0.0025
	11-methylnonacosane	0.1009	0.261	0.0683	0.0308	0.0356	0.5751
	Unknown branched alkane (RI 2953.7)	0.1783	0.6779	0.0003	0.0873	0.3352	0.0209
	11-methyltriacontane	0.0794	0.1512	0.1912	0.06	0.124	0.2093
	Unknown branched alkane (RI 3051.9)	0.1767	0.6931	0.0003	0.093	0.3762	0.0145
	Unknown branched alkane (RI 3126.2)	-0.0014	0.	1.	-0.0205	0.0196	0.7154
Ester	Ethyl 9-hexadecenoate	0.0704	0.1985	0.1222	0.0278	0.0351	0.5751
	Ethyl hexadecanoate	0.0711	0.2025	0.12	0.0278	0.0351	0.5751
	Ethyl (Z,Z)-9,12-Octadecadienoate	0.071	0.1524	0.1911	0.051	0.0525	0.4692
	Ethyl (Z,Z,Z)-9,12,15-Octadecatrienoate	0.1742	0.5202	0.004	0.1049	0.2167	0.0831
	Ethyl 3-trans-Octadecenoate	0.0351	0.0495	0.5188	0.0046	0.001	1.
	(Z)-14-tricosenyl formate	0.0923	0.3127	0.0383	NA	NA	NA
	Oleic acid long ester	0.1086	0.4392	0.0114	0.0509	0.1179	0.2157
	(Z)-Hexenyl hexadecanoate	0.1193	0.4288	0.0125	NA	NA	NA
	Hexyl (E,Z)-9,11-octadecadienoate	0.07	0.1983	0.1222	NA	NA	NA
	Hexyl octadecenoate	0.1491	0.5057	0.0049	NA	NA	NA
	(Z)-3-Hexenyl octadecanoate	0.0774	0.2403	0.0831	NA	NA	NA
Fatty acid	Palmitic acid	0.1144	0.3743	0.0203	0.0074	0.0025	1.
	(Z,Z)-9,12-Octadecadienoic acid	0.0157	0.0087	0.8884	0.0515	0.0503	0.4783
	Stearic acid	0.139	0.4229	0.0129	0.0586	0.102	0.258
	Docosanoic acid	0.0637	0.1572	0.1859	NA	NA	NA
Macrolide /lactone	2-Nonadecanone	-0.0738	0.1582	0.1859	-0.0036	0.0006	1.
	14-Tetradecanolide	0.1217	0.3922	0.018	NA	NA	NA
	16-Hexadecanolide	0.0886	0.2634	0.0683	0.0349	0.0557	0.4531
Nitrogen	(Z)-13-Docosenamide	-0.0175	0.0104	0.8663	0.0162	0.0115	0.8299
Phenol	α-Tocopherol	0.1104	0.4292	0.0125	-0.0219	0.0216	0.6969
	Benzeneacetaldehyde	0.1212	0.3785	0.0203	NA	NA	NA
Sterol	3α,5-Cyclo-5α-cholestan-6-β-ol	0.0925	0.3315	0.0307	NA	NA	NA
	Cholesta-3,5-diene	0.1263	0.3608	0.023	0.0105	0.0037	0.9873
	Cholesterol	0.063	0.0987	0.3027	-0.0216	0.0235	0.6805
Terpene	(E)-Ocimene	0.2503	0.8713	0.	NA	NA	NA
	Alloocimene	0.1141	0.3472	0.0265	NA	NA	NA
	Squalene	0.0086	0.0015	1.	0.0258	0.028	0.6371

B	*H. charithonia*	Male genitals			Female genitals		
Class	Compound	Estimate	R^2	FDR p-value	Estimate	R^2	FDR p-value
Alcohol	2-(cyclohexylmethyl)butane-1,3-diol	0.0776	0.2453	0.0521	NA	NA	NA
	(Z)-9-Octadecen-1-ol	0.0902	0.2946	0.0408	0.1262	0.2914	0.0672
	1-Octadecanol	0.0767	0.2876	0.0413	0.1121	0.3688	0.0413
	1-Icosanol	0.1253	0.3767	0.0181	0.1381	0.5308	0.0102
	1-Docosanol	0.1	0.4224	0.0102	0.1021	0.3526	0.0427
Alkane	Henicosane	0.0823	0.2627	0.0438	0.1687	0.6523	0.0031
	Tricosane	0.0788	0.274	0.0427	0.2369	0.8659	0.
	Pentacosane	0.0682	0.2425	0.0531	0.2017	0.8251	0.0001
	Heptacosane	0.0522	0.153	0.1472	0.1575	0.6955	0.0015
	Nonacosane	0.0592	0.1776	0.1119	0.1889	0.8088	0.0001
	Hentriacontane	0.0274	0.0384	0.5191	0.1182	0.3455	0.0438
Alkene	(Z)-9-Tricosene	0.0993	0.3535	0.0213	0.1653	0.6022	0.0041
	Pentacosene	0.1318	0.4323	0.01	0.1022	0.2614	0.0877
	Heptacosene	0.0888	0.3166	0.033	0.1472	0.3948	0.0366
	Hentriacontene	0.0824	0.2992	0.0396	0.1505	0.4648	0.0181
Branched alkane	11-Methyltricosane	0.1113	0.3683	0.0181	0.1272	0.5114	0.0123
	11-Methylheptacosane	0.1032	0.3742	0.0181	0.0592	0.075	0.4355
	5-Methylheptacosane	0.0807	0.2844	0.0413	0.0975	0.3522	0.0427
	Unknown branched alkane (RI 2859.4)	0.051	0.1224	0.1937	0.063	0.1427	0.233
	11-Methylnonacosane	0.1067	0.3726	0.0181	0.1109	0.3378	0.0443
	5-Methylnonacosane	0.1101	0.4562	0.0071	0.1457	0.6407	0.0033
	Unknown branched alkane (RI 2980.1)	0.0438	0.0964	0.2626	0.0502	0.0484	0.5313
Ester	Ethyl (Z,Z,Z)-9,12,15-Octadecatrienoate	0.0765	0.2741	0.0427	0.0835	0.1729	0.1875
	Ethyl oleate	0.0508	0.1284	0.1875	NA	NA	NA
	Oleyl acetate	0.0993	0.2787	0.0427	0.1936	0.6191	0.0041
	Octadecyl acetate	0.0417	0.0543	0.4356	0.0582	0.1051	0.3244
	(Z)-3-Hexenyl hexadecanoate	0.1631	0.4925	0.0041	NA	NA	NA
	Icosyl acetate	0.0617	0.1381	0.1735	0.1145	0.2171	0.1325
	(Z)-3-Hexenyl octadecenoate	0.1045	0.2869	0.0413	NA	NA	NA
	Hexyl octadecenoate	0.107	0.3491	0.0218	0.0333	0.04	0.5721
Fatty acid	Palmitic acid	0.102	0.3372	0.0251	0.1066	0.336	0.0443
	(Z,Z)-9,12-Octadecadienoic acid	0.0512	0.1353	0.1771	-0.0073	0.0014	1.
	Stearic acid	0.0716	0.2398	0.0539	0.0744	0.1916	0.1636
Nitrogen	Unknown nitrogen compound	0.0706	0.1579	0.1404	0.0049	0.0005	1.
	(Z)-9-Octadecenamide	-0.0177	0.0137	0.7386	0.0415	0.0553	0.5073
	(Z)-13-Docosenamide	-0.0363	0.0505	0.4484	-0.0556	0.1111	0.3091
Phenol	a-Tocopherol	0.0251	0.0319	0.5553	0.0339	0.0278	0.6629
	b-Tocopherol	0.0719	0.2146	0.0733	0.1554	0.3432	0.0438
	c-Tocopherol	0.0299	0.0483	0.4574	0.0299	0.0214	0.7184
Plant	cis-3-Hexenyltiglate	0.0628	0.1841	0.1045	NA	NA	NA
	Neophytadiene	0.0854	0.2071	0.0796	NA	NA	NA
	Benzyl salicylate	0.1348	0.4898	0.0041	0.0796	0.1582	0.2018
	(E)-Phytol	-0.0398	0.0713	0.3506	-0.0556	0.1111	0.3091
Sterol	Cholesta-3,5-diene	0.0036	0.0006	1.	0.0086	0.0017	1.
	(3b)-Cholest-5-en-3-ol	NA	NA	NA	0.0716	0.168	0.1893
	Cholesterol	0.0052	0.0011	1.	-0.0121	0.0041	0.97
Terpene	Squalene	0.0014	0.0001	1.	0.0517	0.0481	0.5313
Unknown	Unknown m-z 83-107-168	0.0383	0.0509	0.4484	NA	NA	NA
	Unknown di-syringic acid compound (RI 2976.8)	0.0492	0.1253	0.1893	0.0961	0.3479	0.0437

Supp. Table 4.3. Linear regression results for all compounds in the genitals of A) *H. atthis* and B) *H. charithonia.* The p-values shown here were corrected according to the Benjamini-Hochberg procedure for controlling for the False Discovery Rate (Benjamini and Hochberg, 1995). The estimate columns indicate the intercept, and R^2 the proportion of variance explained by the linear model. Highlighted cells indicate significant corrected p-values, and compounds that never show correlations with age are shown in grey.

General Discussion

Heliconius research has come a long way since their bright colours first caught the attention of Bates in 1862 (Bates, 1862). Research on their chemical ecology is relatively still at an early stage, but seems more and more promising. However, a macroevolutionary perspective of pheromone differentiation in the Heliconiini was missing, as was chemical ecology data on the non-*Heliconius* Heliconiini altogether. We also had no experimental research on the factors affecting the timing of pheromone production in *Heliconius.* Here, I produced the most inclusive dataset on Heliconiini androconial and genital extracts to date, featuring over half of all described Heliconiini species, and used this new data to test evolutionary hypotheses.

Summary of results

In Chapter 1, I was able to demonstrate that *Heliconius* and *Eueides* androconial extracts were fundamentally different from those of "basal" species. Rather unexpectedly, "basal" Heliconiini genera (*Dryas, Dryadula, Dione, Agraulis* and *Philaethria*) do not produce high levels of fatty acid (FA) derivatives as part of their androconial blends, in contrast with most other Lepidoptera, and in contrast with the species-rich *Heliconius* and *Eueides* genera*. Heliconius* and *Eueides*, by re-upregulating their FA metabolism in the androconia, gained access to a wide range of compounds that was probably key to the fast male sex pheromone (MSP) diversification observed in *Heliconius.* This, most strikingly, allowed for extreme MSP differentiation between sister species pairs in sympatry, a sign of reproductive character displacement and very convincing point in favour of pheromones as important factors in reproductive isolation across the tribe. To further corroborate this, in Chapter 2, I was able to use F2 hybrid phenotypes to map MSP differences between one sympatric sister pair, *Heliconius pardalinus* and *Heliconius elevatus* (Rosser *et al.*, 2015, 2019) and found that most of the signature compounds from either species map to two areas of high genomic divergence, on chromosome 19 and 20, ideal locations for "speciation genes". Within these regions, I found a large number of FA metabolism-related enzymes. Only one of these two regions had previously been reported as containing pheromone QTLs (Byers *et al.*, 2020, 2021) controlling differences between *Heliconius melpomene* and *Heliconius cydno*.

Conversely, in the antiaphrodisiac-centric Chapter 3, I did not find any differences of such a magnitude between the various Heliconiini genera, nor did I find any effect of sympatry, albeit the genital blends remain extremely varied and evolutionarily labile. While antiaphrodisiac signals may be effective at repelling heterospecific males, their action is only effective after mating (Gilbert, 1976; Schulz *et al.*, 2007, 2008), making them ineffective at maintaining reproductive isolation. I also could not detect any effect of mating strategy on pheromone diversification, unlike what reported in (Estrada *et al.*, 2011).

I was able to observe that genital blend composition in many species is comprised of one or few very volatile compounds accompanied by a rich and diverse matrix of esters, as reported in a subset of species in (Estrada *et al.*, 2011), and proposed that these early-eluting volatiles may have the strongest behavioural effect of the blend, taking a cue from the early eluting volatile (E)-β-ocimene in *H. melpomene,* without which the genital blend loses much of its antiaphrodisiac properties (Schulz *et al.*, 2008). I tested this on one of the volatiles from my dataset, 4-hydroxycyclopent-2-en-1-one, but the observation that it was found in both virgin females as well males, makes it not suitable as an antiaphrodisiac. However, I showed that the presence of this volatile was determined by the larval host plant, implying that environmental factors may affect or even facilitate the evolution of diverse pheromone blends. An effect of larval host plant on pheromone composition in *Heliconius* has been reported in the past for MSPs (Darragh, Byers, *et al.*, 2019) but not for genital extracts.

In Chapter 4 I took a departure from broad evolutionary patterns and phylogenetic comparisons to investigate a basic (but still broadly not understood) aspect of *Heliconius* chemical ecology: the timing of pheromone production. I tested the effect of age on the variety of chemical species produced in the free-mating *H. atthis* and the pupal mating *H. charithonia,* and found that the production of most compounds peaks on day 6-8 post-emergence. This coincides with the anecdotally reported onset of sexual maturity in *Heliconius,* and it is one of only a few studies on the effect of aging on their physiological processes (Karlsson, 1994; Boggs, 1997; Finkbeiner, 2014; Montgomery, Merrill and Ott, 2016; Sculfort *et al.*, 2021), and the first on pheromones. As no experiments have ever been carried out on the behavioural activity of these species' compounds, we do not know which components of their androconial and genital blend are truly pheromones, but compounds that are both age-dependent (peaking with the onset of sexual maturity) and exclusive to male androconia or genitals may be good candidates for the role of MSPs and antiaphrodisiacs respectively.

Future prospects: what is missing?

The past decade has seen a growing interest in the involvement of pheromones in the *Heliconius* evolutionary radiation, yet we still know little about their evolution and their functioning. Having researched a broad range of different topics pertaining to them, it has been easy for me to identify areas where is research is limited or lacking. This is especially evident when comparing Heliconiini butterflies to other taxa where chemical ecology has been investigated more in-depth, such as Diptera, Hymenoptera, moths (see for example Raina, Klun and Stadelbacher, 1986; Löfstedt *et al.*, 1989; Rafaeli and Soroker, 1989), or even other butterflies like *Bicyclus* (Nieberding *et al.*, 2012; Karl, Heuskin and Fischer, 2013; Dion, Monteiro and Yew, 2016).

The most obvious information we are still broadly lacking concerns the behavioural effect of most compounds found in this study, and how females interpret the signals. In *H. melpomene,* it has been found that the aldehyde octadecanal is the most behaviourally active compound in the androconial blends, causing strong reactions in female antennae of both *H. melpomene* and its close relative *H. cydno* (Byers *et al.*, 2020). Among the antiaphrodisiac compounds, only (E)-β-ocimene has been functionally tested (Schulz *et al.*, 2008). While in this project I have reported the androconial and genital blends of many species, most of them are lacking such tests and therefore, it is as of yet impossible to determine which compounds in the extract are truly behaviourally active and which are merely structural. As described in Chapter 4, compounds that show a positive correlation with age and peak or plateau with the onset of sexual maturity, while also being exclusive to male tissues (or mated female tissues in the case of antiaphrodisiacs), may be more likely to be behaviourally active in reproduction. Nonetheless, more direct functional tests would be necessary to correctly identify the real pheromone compounds of each species.

Real-time headspace analysis for functional testing. The current method for pheromone extraction in *Heliconius* involves tissue extraction from freshly dead individuals, leaving a lot of uncertainty as to which of the extracted chemicals are actually emitted by the butterflies. Therefore, it would be interesting to design a technique where chemical bouquets may be extracted from live butterflies in real time. One method that may help elucidate which volatile compounds are 1) actually released by live butterflies and 2) released during courtship, is headspace analysis with solid phase microextraction (SPME) (Vas and Vékey, 2004). SPME is a technique that involves a chamber and a syringe-like device that is immersed in it. The syringe is equipped with a silica fibre coated in a thin film of a spongy polymer (this is the "solid phase") that can capture molecules contained in the chamber. Once this capture is complete, the fibre can be transferred to a GCMS instrument for analysis of the collected molecules (Vas and Vékey, 2004). This works in both a gaseous and a liquid environment, and in case of the former, it is referred to as headspace SPME (Vas and Vékey, 2004). This method has been used commonly to analyze plant emissions, but there are examples of it being used to detect insect signals. Headspace SPME would not so much be an alternative to either tissue extractions or behavioural experiments, but rather a complement. Headspace SPME can work on extracted tissues or whole body samples (Rochat *et al.*, 2000; Chu, Hung and Hsu, 2005; Stanley *et al.*, 2018) as well as on living individuals (Rochat *et al.*, 2000; Pontes *et al.*, 2008; Farine, Ferveur and Everaerts, 2012; Silva *et al.*, 2017; Stanley *et al.*, 2018; Rahmani *et al.*, 2020; Zhang *et al.*, 2022), and detected headspace volatiles would take major priority in behavioural and electroantennographic functional assays over other compounds, and furthermore headspace SPME has also been used in the past to detect emissions in real time from individuals during the act of mating (Pontes *et al.*, 2008).

On the other hand, if non-headspace compounds were to also have a behavioural response, it may imply a form of short-range signalling in *Heliconius.* The difficulty of performing headspace SPME on *Heliconius* would mainly lie in the experimental set up, in preparing a chamber large enough to allow for the butterflies to perform their natural behaviours.

Functional assays in a sympatric pair. Somewhat confusingly, when electroantennographic (EAG) assays were conducted (Byers *et al.*, 2020), *H. melpomene* and its sister *H. cydno* both showed strong responses to the compound, while it is only present in the former, implying that the peripheral sensory infrastructure needed to perceive the compound has not diverged between the two species (Byers *et al.*, 2020). This is in contrast to what is expected under a scenario of sensory drive, where divergence of the signal is linked to divergence of the receptors (Endler, 1992). This lack of the necessary sensory sensitivity to perceive the main difference in chemical signature between conspecifics cannot be an indicator that MSPs are not involved in mating, as we know that removing them puts a male at a mating disadvantage (Darragh *et al.*, 2017). On one hand, it is possible that the reaction observed in *H. cydno* antennae may produce a negative response in a living butterfly, a distinction that is impossible to make in EAG assays. On the other hand, in order to observe divergent antennal responses we may have to analyse a pair of species at a different stage of the speciation continuum. *H. melpomene* and *H. cydno* have broadly overlapping ranges (Rosser *et al.*, 2015) and are able to produce F1 and backcross hybrids (Naisbit, Jiggins and Mallet, 2001; Jiggins, 2008; Jiggins *et al.*, 2008; Byers *et al.*, 2020), but they are not considered sister species: the sister species to *H. cydno* is actually the mostly allopatric *H. pachinus* (Kozak *et al.*, 2015; Rosser *et al.*, 2015). A very suitable pair of species to continue testing female responses would be a sympatric sister species pair, making the *H. elevatus* and *H. pardalinus butleri* ideal candidates. Comparatively, these two species diverged a shorter time ago and are still able to produce interfertile F1 hybrids, and so F2 hybrids. It would be interesting to observe the reaction of female antennae to compounds from conspecifics and heterospecifics to see if the responses are more species-dependent in a pair that has not diverged as long ago as *H. melpomene* and *H. cydno,* and that requires strong and focused responses due to the overlapping ranges, as a result of reinforcement.

Other environmental factors affecting pheromone production. On the topic of female responses, the information conveyed by MSPs and how it affects female choice is still very poorly understood. In both Chapter 3 and in (Darragh, Byers, *et al.*, 2019) it has been shown that larval host plant can alter the chemical profiles of species, and I have found an effect of aging on most androconial and genital compounds in Chapter 4, so male host plant and age are two pieces of information that the pheromones may carry. The Chapter 4 age effect study was originally designed to include tests for other factors as well, such as time of the day, presence/absence of females, courting vs. non-courting

(in this case the androconia and genitals would be extracted in real time from courting and resting males), feeding status. Time of the day in particular is known to be important to pheromone emissions in Lepidoptera (Baker and Cardé, 1979; Raina, Klun and Stadelbacher, 1986; Rafaeli and Soroker, 1989). Testing the effect of these factors would essentially decode what kinds of information on the male may be carried by the volatiles, which would have multiple implications: it would potentially reveal some regulatory information on pheromone release, some factors affecting female choice, and also produce data that would facilitate future functional and behavioural assays. Given my results in Chapter 3 and 4, it is especially evident that adult male androconial and genital composition may be very variable depending on whether the butterflies utilized in the study were wild-caught or captive, and on their housing and rearing conditions. Knowing what affects pheromone composition may help mitigate the effect of anthropogenically introduced variables.

In-depth genetic analysis. I identified a number of quantitative trait loci linked to differences in chemical signature between the sympatric sister species *H. elevatus* and *H. pardalinus butleri*, and more pheromone QTLs were discovered in the closely related *H. melpomene* and *H. cydno* (Darragh, Orteu, *et al.*, 2019; Byers *et al.*, 2020, 2021). In my study, QTL mapping was performed on F2 hybrids, whereas backcrosses were used in the *H. melpomene-H. cydno* studies due to the inability of these two species to produce F2 hybrids. However, data on backcrosses is available for *H. elevatus* and *H. pardalinus,* so the next step in QTL mapping for that species pair should be to incorporate backcross individuals. Following this, it would be important to search for epistatic interactions between loci that may be the cause for some of the contradictory phenotype:genotype effects observed in F2 hybrids (where high concentrations of a *H. elevatus* compound appear linked to homozygous *H. pardalinus* alleles and vice versa).

Some of the genes encoded within the *H. elevatus/H. pardalinus* QTLs are present in a high number of copies and are rather generic in function, while being surely linked to fatty acid metabolism, like fatty acyl-CoA reductases (Reed *et al.*, 1994, 1995; Tillman *et al.*, 1999; Liénard *et al.*, 2014; He *et al.*, 2017; Mann *et al.*, 2017), and may potentially have redundancy with other loci in the genome. However, loci on chromosome 19 that include fatty acyl-CoA Δ9-desaturases may be good targets for functional testing with CRISPR knock-outs. CRISPR has been used to produce *Heliconius* colour pattern mutants, and the result are mosaic individuals (Mazo-Vargas *et al.*, 2017; Westerman *et al.*, 2018; Concha *et al.*, 2019; McMillan *et al.*, 2020; Livraghi *et al.*, 2021). The effects mosaic knockouts on pheromone composition may not be as clear to see as they are on the colour pattern, but one may still be able to observe them. For example, *H. elevatus* alkanes map to loci for Δ9-desaturases; as Δ9-desaturases are involved in the production of *H. pardalinus* compounds, alkane production may be linked to their deactivation. Thus, targeting Δ9-desaturases would potentially lead to alkane biosynthesis.

Site of pheromone production. Part of my past research involved unreported deuterium-labelling experiments aimed at validating the fatty acid metabolic pathway reported by (Mann *et al.*, 2017) *in vivo.* These experiments involve supplying butterfly tissues with a deuterium-labelled precursor (e.g. palmitic acid) via topical application on the androconia, and then observing deuterium incorporation into intermediate compounds. This method worked on *Bicyclus anynana* (Liénard *et al.*, 2014; Wang *et al.*, 2014)*,* but not in *Heliconius.* Alternatively, deuterium-labelled precursors can be fed to the larvae (Schulz, Yildizhan and van Loon, 2011), but this also yielded no results in *Heliconius.* There are many reasons why this may not have worked, including wrong timing, extremely small sample size, or wrong site of topical application, among others. But with a better experimental design, and a more thorough knowledge of factors affecting pheromone biosynthesis, these experiments may yield results.

Concluding remarks

Chemical ecology is a fascinating topic not only in *Heliconius* and in the Heliconiini tribe, but also across the whole tree of life. It is exciting to research such an ancient and widespread form of communication, and work in this field can be very informative of a clade's evolutionary history and speciation processes. In my research, I have had the opportunity of exploring such broad scope topics as macroevolution of chemical signals and how they may impact reproductive isolation, and how different types of signals within the same clade may evolve in radically different directions. I was able to explore the genomic mechanisms controlling putative pheromone divergence in sister species. Even more basic topics, such as the effect of age on chemical signature, proved very informative. There is still much to learn about pheromones in *Heliconius* and in speciation in general, yet I trust that my work has not only answered some big questions, but also laid the foundation for future work. With pheromones appearing a more complex and multifaceted topic than ever, it will be exciting to see future developments in this field of research.

References

Ablard, K. M. *et al.* (2014) 'First-male sperm precedence and precopulatory and postcopulatory rituals in the parasitoid wasp *Ooencyrtus kuvanae* (Hymenoptera: Encyrtidae)', *Canadian Entomologist*. Cambridge University Press, 146(5), pp. 548–559. doi: 10.4039/tce.2014.7.

Adams, D. C. (2014) 'A generalized K statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data', *Systematic Biology*. Oxford University Press, 63(5), pp. 685–697. doi: 10.1093/sysbio/syu030.

Adams, D. C. and Collyer, M. L. (2018) 'Multivariate Phylogenetic Comparative Methods: Evaluations, Comparisons, and Recommendations', *Systematic Biology*. Oxford University Press, 67(1), pp. 14–31. doi: 10.1093/sysbio/syx055.

Adams, D. C. and Otárola-Castillo, E. (2013) 'Geomorph: An r package for the collection and analysis of geometric morphometric shape data', *Methods in Ecology and Evolution*, 4(4), pp. 393–399. doi: 10.1111/2041-210X.12035.

Albre, J., Steinwender, B. and Newcomb, R. D. (2013) 'The Evolution of Desaturase Gene Regulation Involved in Sex Pheromone Production in Leafroller Moths of the Genus Planotortrix', *Journal of Heredity*. Oxford Academic, 104(5), pp. 627–638. doi: 10.1093/jhered/est043.

Alfaro, M. E. *et al.* (2009) 'Does evolutionary innovation in pharyngeal jaws lead to rapid lineage diversification in labrid fishes?', *BMC Evolutionary Biology*. BioMed Central, 9(1), p. 255. doi: 10.1186/1471-2148-9-255.

Allison, J. D. and Carde, R. T. (2017) *Pheromone Communication in Moths: Evolution, Behavior, and Application*. Oakland: University of California Press. doi: 10.1093/ae/tmx069.

Anderson, M. J. (2006) 'Distance-based tests for homogeneity of multivariate dispersions', *Biometrics*. Biometrics, 62(1), pp. 245–253. doi: 10.1111/j.1541-0420.2005.00440.x.

Anderson, M. J., Ellingsen, K. E. and McArdle, B. H. (2006) 'Multivariate dispersion as a measure of beta diversity', *Ecology Letters*. John Wiley & Sons, Ltd, 9(6), pp. 683–693. doi: 10.1111/j.1461-0248.2006.00926.x.

Andersson, J. *et al.* (2007) 'Male sex pheromone release and female mate choice in a butterfly.', *The Journal of experimental biology*, 210(Pt 6), pp. 964–70. doi: 10.1242/jeb.02726.

Andersson, J., Borg-Karlson, A. K. and Wiklund, C. (2000) 'Sexual cooperation and conflict in butterflies: a male-transferred anti-aphrodisiac reduces harassment of recently mated females', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 267(1450), pp. 1271–1275. doi: 10.1098/rspb.2000.1138.

Andersson, J., Borg-Karlson, A. K. and Wiklund, C. (2003) 'Antiaphrodisiacs in pierid butterflies: A theme with variation!', *Journal of Chemical Ecology*, 29(6), pp. 1489–1499. doi: 10.1023/A:1024277823101.

Andersson, J., Borg-Karlson, A. K. and Wiklund, C. (2004) 'Sexual conflict and anti-aphrodisiac titre in a polyandrous butterfly: Male ejaculate tailoring and absence of female control', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 271(1550), pp. 1765–1770. doi: 10.1098/rspb.2003.2671.

Ando, T., Inomata, S. and Yamamoto, M. (2004) 'Lepidopteran Sex Pheromones', in. Springer, Berlin,

Heidelberg, pp. 51–96. doi: 10.1007/b95449.

Andriamaharavo, N. R. (2014) 'Retention Data. NIST Mass Spectrometry Data Center.'

Arak, A. and Enquist, M. (1995) 'Conflict, receiver bias and the evolution of signal form', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, 349(1330), pp. 337–344. doi: 10.1098/rstb.1995.0122.

Arnegard, M. E. *et al.* (2010) 'Sexual Signal Evolution Outpaces Ecological Divergence during Electric Fish Species Radiation'. doi: 10.1086/655221.

Arnqvist, G. and Nilsson, T. (2000) 'The evolution of polyandry: Multiple mating and female fitness in insects', *Animal Behaviour*. Academic Press, pp. 145–164. doi: 10.1006/anbe.2000.1446.

Arnqvist, G. and Rowe, L. (1995) 'Sexual conflict and arms races between the sexes: A morphological adaptation for control of mating in a female insect', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 261(1360), pp. 123–127. doi: 10.1098/rspb.1995.0126.

Arnqvist, G. and Rowe, L. (2005) *Sexual conflict*. Princeton: Princeton University Press.

Arrese, E. L. and Soulages, J. L. (2010) 'Insect fat body: Energy, metabolism, and regulation', *Annual Review of Entomology*. NIH Public Access, pp. 207–225. doi: 10.1146/annurev-ento-112408-085356.

Ayasse, M., Paxton, R. J. and Tengö, J. (2001) 'Mating behavior and chemical communication in the order Hymenoptera', *Annual Review of Entomology*, pp. 31–78. doi: 10.1146/annurev.ento.46.1.31.

Ayre, G. L. and Blum, M. S. (1971) 'Attraction and Alarm of Ants (*Camponotus* spp.: Hymenoptera: Formicidae) by Pheromones', *Physiological Zoology*. University of Chicago Press, 44(2), pp. 77–83. doi: 10.1086/physzool.44.2.30155558.

Bacquet, P. M. B. *et al.* (2015) 'Selection on male sex pheromone composition contributes to butterfly reproductive isolation', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society of London, 282(1804), p. 20142734. doi: 10.1098/rspb.2014.2734.

Baird, N. A. *et al.* (2008) 'Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers', *PLoS ONE*. Edited by J. C. Fay. Public Library of Science, 3(10), p. e3376. doi: 10.1371/journal.pone.0003376.

Baker, T. C. and Cardé, R. T. (1979) 'Endogenous and exogenous factors affecting periodicities of female calling and male sex pheromone response in *Grapholitha molesta* (Busck)', *Journal of Insect Physiology*. Pergamon, 25(12), pp. 943–950. doi: 10.1016/0022-1910(79)90107-0.

Barbash, D. A. *et al.* (2003) 'A rapidly evolving MYB-related protein causes species isolation in *Drosophila*', *Proceedings of the National Academy of Sciences of the United States of America*, 100(9), pp. 5302–5307. doi: 10.1073/pnas.0836927100.

Barbash, D. A., Awadalla, P. and Tarone, A. M. (2004) 'Functional Divergence Caused by Ancient Positive Selection of a *Drosophila* Hybrid Incompatibility Locus', *PLoS Biology*. Edited by Mohamed Noor. Public Library of Science, 2(6), p. e142. doi: 10.1371/journal.pbio.0020142.

Bates, H. W. (1862) 'XXXII. Contributions to an Insect Fauna of the Amazon Valley. Lepidoptera: Heliconidae.', *Transactions of the Linnean Society of London*. Blackwell Publishing Ltd, 23(3), pp. 495–566. doi: 10.1111/j.1096-3642.1860.tb00146.x.

Baxter, S. W. *et al.* (2008) 'Convergent Evolution in the Genetic Basis of Müllerian Mimicry in *Heliconius* Butterflies', *Genetics*, 180(3), pp. 1567–77.

Van Belleghem, S. M. *et al.* (2017) 'Complex modular architecture around a simple toolkit of wing

pattern genes', *Nature Ecology and Evolution*. Nature Publishing Group, 1(3). doi: 10.1038/s41559-016-0052.

Bellwood, D. R. *et al.* (2010) 'Evolutionary history of the butterflyfishes (f: Chaetodontidae) and the rise of coral feeding fishes', *Journal of Evolutionary Biology*. John Wiley & Sons, Ltd (10.1111), 23(2), pp. 335–349. doi: 10.1111/j.1420-9101.2009.01904.x.

Beltrán, D. F., Shultz, A. J. and Parra, J. L. (2021) 'Speciation rates are positively correlated with the rate of plumage color evolution in hummingbirds', *Evolution*, p. evo.14277. doi: 10.1111/evo.14277.

Beltrán, M. *et al.* (2007) 'Do pollen feeding, pupal-mating and larval gregariousness have a single origin in *Heliconius* butterflies? Inferences from multilocus DNA sequence data', *Biological Journal of the Linnean Society*. Oxford Academic, 92(2), pp. 221–239. doi: 10.1111/j.1095-8312.2007.00830.x.

Bengtsson, M. *et al.* (2014) 'Pheromone races of *Cydia splendana* (Lepidoptera, Tortricidae) overlap in host plant association and geographic distribution', *Frontiers in Ecology and Evolution*. Frontiers Media S. A, 2(AUG), p. 46. doi: 10.3389/fevo.2014.00046.

Benjamini, Y. and Hochberg, Y. (1995) 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing', *Journal of the Royal Statistical Society: Series B (Methodological)*. Wiley, 57(1), pp. 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x.

Birkhead, T. R. (1998) 'Sperm competition in birds', *Reviews of Reproduction*. Rev Reprod, 3(2), pp. 123–129. doi: 10.1530/ror.0.0030123.

Bjostad, L. B. and Roelofs, W. L. (1981) 'Sex pheromone biosynthesis from radiolabeled fatty acids in the redbanded leafroller moth.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 256(15), pp. 7936–40.

Blair, W. F. (1974) 'Character Displacement in Frogs', *American Zoologist*. Oxford Academic, 14(4), pp. 1119–1125. doi: 10.1093/icb/14.4.1119.

Blomberg, S. P., Garland, T. and Ives, A. R. (2003) 'Testing for phylogenetic signal in comparative data: behavioral traits are more labile', *Evolution*. John Wiley & Sons, Ltd, 57(4), pp. 717–745. doi: 10.1111/j.0014-3820.2003.tb00285.x.

Blomquist, G. J., Dillwith, J. W. and Adams, T. S. (1987) 'Biosynthesis and Endocrine Regulation of Sex Pheromone Production in Diptera', in *Pheromone Biochemistry*. Elsevier, pp. 217–250. doi: 10.1016/b978-0-12-564485-3.50012-9.

Boch, R. and Rothenbuhler, W. C. (1974) 'Defensive behaviour and production of alarm pheromone in honeybees', *Journal of Apicultural Research*. Taylor & Francis, 13(4), pp. 217–221. doi: 10.1080/00218839.1974.11099783.

Boggs, C. L. (1981) 'Nutritional and Life-History Determinants of Resource Allocation in Holometabolous Insects', *The American Naturalist*. University of Chicago Press, 117(5), pp. 692–709. doi: 10.1086/283753.

Boggs, C. L. (1997) 'Reproductive allocation from reserves and income in butterfly species with differing adult diets', *Ecology*. Ecological Society of America, 78(1), pp. 181–191. doi: 10.1890/0012-9658(1997)078[0181:RAFRAI]2.0.CO;2.

Bolnick, D. I. and Fitzpatrick, B. M. (2007) 'Sympatric Speciation: Models and Empirical Evidence', *Annual Review of Ecology, Evolution, and Systematics*. Annual Reviews, 38(1), pp. 459–487. doi: 10.1146/annurev.ecolsys.38.091206.095804.

Bontonou, G., Denis, B. and Wicker-Thomas, C. (2013) 'Interaction between temperature and male pheromone in sexual isolation in *Drosophila melanogaster*', *Journal of Evolutionary Biology*. J Evol

Biol, 26(9), pp. 2008–2020. doi: 10.1111/jeb.12206.

Boughman, J. W. (2001) 'Divergent sexual selection enhances reproductive isolation in sticklebacks', *Nature*. Nature Publishing Group, 411(6840), pp. 944–948. doi: 10.1038/35082064.

Boughman, J. W. (2002) 'How sensory drive can promote speciation', *Trends in Ecology and Evolution*. Elsevier, pp. 571–577. doi: 10.1016/S0169-5347(02)02595-8.

Boul, K. E. *et al.* (2007) 'Sexual selection drives speciation in an Amazonian frog', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 274(1608), pp. 399–406. doi: 10.1098/rspb.2006.3736.

Bradshaw, J. W. S., Baker, R. and Howse, P. E. (1975) 'Multicomponent alarm pheromones of the weaver ant', *Nature*. Nature Publishing Group, 258(5532), pp. 230–231. doi: 10.1038/258230a0.

Brent, C. S. and Byers, J. A. (2011) 'Female attractiveness modulated by a male-derived antiaphrodisiac pheromone in a plant bug', *Animal Behaviour*, 82(5), pp. 937–943. doi: 10.1016/j.anbehav.2011.08.010.

Van Den Brink, D. M. *et al.* (2005) 'Characterization of the final step in the conversion of phytol into phytanic acid', *Journal of Biological Chemistry*, 280(29), pp. 26838–26844. doi: 10.1074/jbc.M501861200.

Broman, K. W. *et al.* (2003) 'R/qtl: QTL mapping in experimental crosses', *Bioinformatics*. Bioinformatics, 19(7), pp. 889–890. doi: 10.1093/bioinformatics/btg112.

Brown, K. S. (1981) 'The Biology of *Heliconius* and Related Genera', *Annual Review of Entomology*. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, 26(1), pp. 427–457. doi: 10.1146/annurev.en.26.010181.002235.

Brown, W. L. and Wilson, E. O. (1956) 'Character Displacement', *Systematic Zoology*, 5(2), p. 49. doi: 10.2307/2411924.

Brunetti, A. E. *et al.* (2015) 'Frog Volatile Compounds: Application of in vivo SPME for the Characterization of the Odorous Secretions from Two Species of Hypsiboas Treefrogs', *Journal of Chemical Ecology*. Springer New York LLC, 41(4), pp. 360–372. doi: 10.1007/s10886-015-0564-z.

De Bruyne, M. and Baker, T. C. (2008) 'Odor Detection in Insects: Volatile Codes', *J Chem Ecol*, 34, pp. 882–897. doi: 10.1007/s10886-008-9485-4.

Buchinger, T. J. *et al.* (2017) 'Phylogenetic distribution of a male pheromone that may exploit a nonsexual preference in lampreys', *Journal of Evolutionary Biology*. Blackwell Publishing Ltd, 30(12), pp. 2244–2254. doi: 10.1111/jeb.13191.

Butenandt, A., Beckmann, R. and Hecker, E. (1961) 'Über den Sexuallockstoff des Seidenspinners, I: Der biologische Test und die Isolierung des reinen Sexuallockstoffes Bombykol', *Hoppe-Seyler's Zeitschrift fur Physiologische Chemie*, 324(1), pp. 71–83. doi: 10.1515/bchm2.1961.324.1.71.

Byers, K. J. R. P. *et al.* (2020) 'A major locus controls a biologically active pheromone component in *Heliconius melpomene*', *Evolution*. Society for the Study of Evolution, 74(2), pp. 349–364. doi: 10.1111/evo.13922.

Byers, K. J. R. P. *et al.* (2021) 'Clustering of loci controlling species differences in male chemical bouquets of sympatric *Heliconius* butterflies', *Ecology and Evolution*. John Wiley and Sons Ltd, 11(1), pp. 89–107. doi: 10.1002/ece3.6947.

De Cáceres, M. and Legendre, P. (2009) 'Associations between species and groups of sites: Indices and statistical inference', *Ecology*. Ecology, 90(12), pp. 3566–3574. doi: 10.1890/08-1823.1.

De Cáceres, M., Legendre, P. and Moretti, M. (2010) 'Improving indicator species analysis by combining groups of sites', *Oikos*. Blackwell Publishing Ltd, 119(10), pp. 1674–1684. doi: 10.1111/j.1600-0706.2010.18334.x.

Cama, B. (2016) 'Genetic analysis of wing pattern and pheromone composition in two sister species of *Heliconius* butterflies.' University of York.

Cardé, R. T. *et al.* (1977) 'Sex pheromone specificity as a reproductive isolating mechanism among the sibling species Archips argyrospilus and A. mortuanus and other sympatric tortricine moths (Lepidoptera: Tortricidae)', *Journal of Chemical Ecology*. Kluwer Academic Publishers-Plenum Publishers, 3(1), pp. 71–84. doi: 10.1007/BF00988135.

Cardé, R. T. (1986) 'The Role of Pheromones in Reproductive Isolation and Speciation of Insects', in *Evolutionary Genetics of Invertebrate Behavior*. Boston, MA: Springer US, pp. 303–317. doi: 10.1007/978-1-4899-3487-1_29.

Carde, R. T. and Minks, A. K. (1995) 'Control of moth pests by mating disruption: Successes and constraints', *Annual Review of Entomology*. Annual Reviews Inc., pp. 559–585. doi: 10.1146/annurev.en.40.010195.003015.

Carlson, D. A. *et al.* (1971) 'Sex attractant pheromone of the house fly: Isolation, identification and synthesis', *Science*. American Association for the Advancement of Science, 174(4004), pp. 76–78. doi: 10.1126/science.174.4004.76.

Casselton, L. A. (2002) 'Mate recognition in fungi', in *Heredity*. Blackwell Publishing Ltd., pp. 142–147. doi: 10.1038/sj.hdy.6800035.

Casselton, L. A. and Olesnicky, N. S. (1998) 'Molecular Genetics of Mating Recognition in Basidiomycete Fungi', *Microbiology and Molecular Biology Reviews*. American Society for Microbiology, 62(1), pp. 55–70. doi: 10.1128/mmbr.62.1.55-70.1998.

Challi, R. *et al.* (2016) 'Lepbase: the Lepidopteran genome database', *bioRxiv*, p. 056994. doi: 10.1101/056994.

Chamero, P. *et al.* (2007) 'Identification of protein pheromones that promote aggressive behaviour', *Nature*. Nature Publishing Group, 450(7171), pp. 899–902. doi: 10.1038/nature05997.

Chapman, T. *et al.* (2003) 'Sexual conflict', *Trends in Ecology and Evolution*. Elsevier, pp. 41–47. doi: 10.1016/S0169-5347(02)00004-6.

Chaturvedi, R., Singh, M. and Sane, P. V. (1990) 'Photoinactivation of Photosynthetic Electron Transport under Anaerobic and Aerobic Conditions in Isolated Thylakoids of Spinach', in *Current Research in Photosynthesis*. Springer Netherlands, pp. 1353–1356. doi: 10.1007/978-94-009-0511-5_312.

Chu, T. Y., Hung, C. C. and Hsu, C. Y. (2005) 'Solid-phase microextraction for the investigation of sex pheromone of Eucosma notanthes Meyrick', *Talanta*. Elsevier B.V., 65(3), pp. 743–749. doi: 10.1016/j.talanta.2004.08.023.

Chung, H. and Carroll, S. B. (2015) 'Wax, sex and the origin of species: Dual roles of insect cuticular hydrocarbons in adaptation and mating', *BioEssays*. John Wiley and Sons Inc., 37(7), pp. 822–830. doi: 10.1002/bies.201500014.

Cimino, G. *et al.* (1991) 'Alarm pheromones from the Mediterranean opisthobranch *Haminoea navicula*', *Experientia*. Birkhäuser-Verlag, 47(1), pp. 61–63. doi: 10.1007/BF02041253.

Clavel, J., Escarguel, G. and Merceron, G. (2015) 'mvMORPH: An R package for fitting multivariate evolutionary models to morphometric data', *Methods in Ecology and Evolution*. Edited by T. Poisot,

6(11), pp. 1311–1319. doi: 10.1111/2041-210X.12420.

Collignon, R. M. *et al.* (2016) 'The Influence of Host Plant Volatiles on the Attraction of Longhorn Beetles to Pheromones', *Journal of Chemical Ecology*. Springer New York LLC, 42(3), pp. 215–229. doi: 10.1007/s10886-016-0679-x.

Concha, C. *et al.* (2019) 'Interplay between Developmental Flexibility and Determinism in the Evolution of Mimetic *Heliconius* Wing Patterns', *Current Biology*. Cell Press, 29(23), pp. 3996-4009.e4. doi: 10.1016/j.cub.2019.10.010.

Coppée, A. *et al.* (2011) 'Age-dependent attractivity of males' sexual pheromones in *Bombus terrestris* (L.) [Hymenoptera, Apidae]', *Chemoecology*, 21(2), pp. 75–82. doi: 10.1007/s00049-011-0070-x.

Couzin, I. D. and Franks, N. R. (2003) 'Self-organized lane formation and optimized traffic flow in army ants', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 270(1511), pp. 139–146. doi: 10.1098/rspb.2002.2210.

Coyne, J. A. and Orr, H. A. (1989) 'Patterns of speciation in *Drosophila*', *Evolution*. Wiley, 43(2), pp. 362–381. doi: 10.1111/j.1558-5646.1989.tb04233.x.

Coyne, J. A. and Orr, H. A. (2004) *Speciation*. Sunderland: Sinauer Associates, Inc.

Craig, G. B. (1967) 'Mosquitoes: Female monogamy induced by male accessory gland substance', *Science*. Science, 156(3781), pp. 1499–1501. doi: 10.1126/science.156.3781.1499.

Cuvillier-Hot, V. *et al.* (2001) 'Sex, age and ovarian activity affect cuticular hydrocarbons in *Diacamma ceylonense*, a queenless ant', *Journal of Insect Physiology*. Pergamon, 47(4–5), pp. 485–493. doi: 10.1016/S0022-1910(00)00137-2.

Daly, M. (1978) 'The Cost of Mating', *The American Naturalist*. University of Chicago Press, 112(986), pp. 771–774. doi: 10.1086/283319.

Darragh, K. *et al.* (2017) 'Male sex pheromone components in *Heliconius* butterflies released by the androconia affect female choice', *PeerJ*. PeerJ Inc., 5, p. e3953. doi: 10.7717/peerj.3953.

Darragh, K., Orteu, A., *et al.* (2019) 'A novel terpene synthase produces an anti-aphrodisiac pheromone in the butterfly *Heliconius melpomene*', *bioRxiv*. Cold Spring Harbor Laboratory, (September), p. 779678. doi: 10.1101/779678.

Darragh, K., Byers, K. J. R. P., *et al.* (2019) 'Male pheromone composition depends on larval but not adult diet in *Heliconius melpomene*', *Ecological Entomology*. Blackwell Publishing Ltd, 44(3), pp. 397–405. doi: 10.1111/een.12716.

Darragh, K. *et al.* (2020) 'Species specificity and intraspecific variation in the chemical profiles of *Heliconius* butterflies across a large geographic range', *Ecology and Evolution*. John Wiley and Sons Ltd, 10(9), pp. 3895–3918. doi: 10.1002/ece3.6079.

Dasmahapatra, K. K. *et al.* (2012) 'Butterfly genome reveals promiscuous exchange of mimicry adaptations among species', *Nature*, 487(7405), pp. 94–98. doi: 10.1038/nature11041.

Davey, J. W. *et al.* (2016) 'Major improvements to the *Heliconius melpomene* genome assembly used to confirm 10 chromosome fusion events in 6 million years of butterfly evolution', *G3: Genes, Genomes, Genetics*. Genetics Society of America, 6(3), pp. 695–708. doi: 10.1534/g3.115.023655.

Davies, M. J. (2008) 'Chemical communication in the European otter, *Lutra lutra*'. Department of Biological Sciences, The University of Hull.

Davis-Foster, D. (2019) *PyMassSpec — PyMassSpec 2.2.21 documentation*. Available at: https://pymassspec.readthedocs.io/en/master/ (Accessed: 17 June 2020).

Deinert, E. I., Longino, J. T. and Gilbert, L. E. (1994) 'Mate competition in butterflies', *Nature*, pp. 23–24. doi: 10.1038/370023a0.

Dieckmann, U. and Doebeli, M. (1999) 'On the origin of species by sympatric speciation', *Nature*. Nature Publishing Group, 400(6742), pp. 354–357. doi: 10.1038/22521.

Dion, E., Monteiro, A. and Yew, J. Y. (2016) 'Phenotypic plasticity in sex pheromone production in *Bicyclus anynana* butterflies', *Scientific Reports*. Nature Publishing Group, 6(1), p. 39002. doi: 10.1038/srep39002.

Dixon, P. (2003) 'VEGAN, a package of R functions for community ecology', *Journal of Vegetation Science*. John Wiley & Sons, Ltd (10.1111), 14(6), pp. 927–930. doi: 10.1111/j.1654-1103.2003.tb02228.x.

Dong, C. *et al.* (2020) 'Convergent evolution of small molecule pheromones in *Pristionchus* nematodes', *eLife*. eLife Sciences Publications Ltd, 9. doi: 10.7554/eLife.55687.

Drickamer, L. C. and Brown, P. L. (1998) 'Age-related changes in odor preferences by house mice living in seminatural enclosures', *Journal of Chemical Ecology*. Kluwer Academic/Plenum Publishers, 24(11), pp. 1745–1756. doi: 10.1023/A:1022395130489.

Dufrêne, M. and Legendre, P. (1997) 'Species assemblages and indicator species: The need for a flexible asymmetrical approach', *Ecological Monographs*. Wiley, 67(3), pp. 345–366. doi: 10.2307/2963459.

Dulac, C. and Torello, A. T. (2003) 'Molecular detection of pheromone signals in mammals: From genes to behaviour', *Nature Reviews Neuroscience*, 4(7), pp. 551–562. doi: 10.1038/nrn1140.

Dumont, E. R. *et al.* (2012) 'Morphological innovation, diversification and invasion of a new adaptive zone', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 279(1734), pp. 1797–1805. doi: 10.1098/rspb.2011.2005.

Ehlers, S. *et al.* (2021) 'Identification and Composition of Clasper Scent Gland Components of the Butterfly *Heliconius erato* and Its Relation to Mimicry', *ChemBioChem*. John Wiley and Sons Inc, 22(23), pp. 3300–3313. doi: 10.1002/cbic.202100372.

El-Sayed, A. M. *et al.* (2003) 'Geographic variation in pheromone chemistry, antennal electrophysiology, and pheromone-mediated trap catch of North American populations of the obliquebanded leafroller', *Environmental Entomology*. Entomological Society of America, 32(3), pp. 470–476. doi: 10.1603/0046-225X-32.3.470.

Eller, F. and Palmquist, D. (2014) 'Factors Affecting Pheromone Production by the Pepper Weevil, *Anthonomus eugenii* Cano (Coleoptera: Curculionidae) and Collection Efficiency', *Insects*. MDPI AG, 5(4), pp. 909–920. doi: 10.3390/insects5040909.

Eltringham, H. (1925) 'VI. On the Abdominal Glands in *Heliconius* (Lepidoptera).', *Transactions of the Royal Entomological Society of London*, 73(1–2), pp. 269–275. doi: 10.1111/j.1365-2311.1925.tb02862.x.

Eltz, T., Ayasse, M. and Lunau, K. (2006) 'Species-specific antennal responses to tibial fragrances by male orchid bees', *Journal of Chemical Ecology*. Springer, 32(1), pp. 71–79. doi: 10.1007/s10886-006-9352-0.

Emsley, M. G. (1963) 'A morphological study of imagine Heliconiinae (Lep.: Nymphalidae) with a consideration of the evolutionary relationships within the group', *Zoologica : scientific contributions*

of the New York Zoological Society., 48(8), pp. 85--130.

Emsley, M. G. (1965) 'Speciation in *Heliconius* (Lep., Nymphalidae): Morphology and geographic distribution', *Zoologica : scientific contributions of the New York Zoological Society.* Smithsonian Institution, 50(14), pp. 191–254. doi: 10.5962/p.203277.

Endler, J. A. (1992) 'Signals, signal conditions, and the direction of evolution', *American Naturalist*. University of Chicago Press, 139(Suppl.). doi: 10.1086/285308.

Estrada, C. *et al.* (2010) 'Sex-specific chemical cues from immatures facilitate the evolution of mate guarding in *Heliconius* butterflies', *Proceedings of the Royal Society B: Biological Sciences*, 277(1680), pp. 407–413. doi: 10.1098/rspb.2009.1476.

Estrada, C. *et al.* (2011) 'Sexual selection drives the evolution of antiaphrodisiac pheromones in butterflies', *Evolution*. Blackwell Publishing Inc, 65(10), pp. 2843–2854. doi: 10.1111/j.1558-5646.2011.01352.x.

Estrada, C. and Jiggins, C. D. (2008) 'Interspecific sexual attraction because of convergence in warning colouration: Is there a conflict between natural and sexual selection in mimetic species?', *Journal of Evolutionary Biology*. John Wiley & Sons, Ltd, 21(3), pp. 749–760. doi: 10.1111/j.1420-9101.2008.01517.x.

Everaerts, C. *et al.* (2010) '*Drosophila* Cuticular Hydrocarbons Revisited: Mating Status Alters Cuticular Profiles', *PLoS ONE*. Edited by F. Marion-Poll. Public Library of Science, 5(3), p. e9607. doi: 10.1371/journal.pone.0009607.

Farine, J.-P., Ferveur, J.-F. and Everaerts, C. (2012) 'Volatile *Drosophila* Cuticular Pheromones Are Affected by Social but Not Sexual Experience', *PLoS ONE*. Edited by P. Callaerts. Public Library of Science, 7(7), p. e40396. doi: 10.1371/journal.pone.0040396.

Feder, J. L. *et al.* (2003) 'Evidence for Inversion Polymorphism Related to Sympatric Host Race Formation in the Apple Maggot Fly, *Rhagoletis pomonella*', *Genetics*. Oxford Academic, 163(3), pp. 939–953. doi: 10.1093/genetics/163.3.939.

Feder, J. L. *et al.* (2012) 'Establishment of new mutations under divergence and genome hitchhiking.', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. The Royal Society, 367(1587), pp. 461–74. doi: 10.1098/rstb.2011.0256.

Feder, J. L. *et al.* (2013) 'Geographic Mode of Speciation and Genomic Divergence', *Annual Review of Ecology, Evolution, and Systematics*. Annual Reviews Inc., 44(1), pp. 73–97. doi: 10.1146/annurev-ecolsys-110512-135825.

Feder, J. L., Egan, S. P. and Nosil, P. (2012) 'The genomics of speciation-with-gene-flow', *Trends in Genetics*, 28(7), pp. 342–350. doi: 10.1016/j.tig.2012.03.009.

Fedina, T. Y. *et al.* (2012) 'Dietary Effects on Cuticular Hydrocarbons and Sexual Attractiveness in *Drosophila*', *PLoS ONE*. Edited by W. J. Etges. Public Library of Science, 7(12), p. e49799. doi: 10.1371/journal.pone.0049799.

Ferveur, J. F. (2005) 'Cuticular hydrocarbons: Their evolution and roles in *Drosophila* pheromonal communication', *Behavior Genetics*. Springer, pp. 279–295. doi: 10.1007/s10519-005-3220-5.

Fields, S. (1990) 'Phermone response in yeast', *Trends in Biochemical Sciences*. Trends Biochem Sci, 15(7), pp. 270–273. doi: 10.1016/0968-0004(90)90052-D.

Finkbeiner, S. D. (2014) 'Communal roosting in *Heliconius* butterflies (Nymphalidae): roost recruitment, establishment, fidelity, and resource use trends based on age and sex', *Journal of the Lepidopterists' Society*. Univ Calif Irvine, Dept Ecol & Evolutionary Biol, Irvine, CA 92697 USA, 68(1),

pp. 10-16 WE-Science Citation Index Expanded (SCI-EX.

Fitzpatrick, B. M., Fordyce, J. A. and Gavrilets, S. (2008) 'What, if anything, is sympatric speciation?', *Journal of evolutionary biology*, 21(6), pp. 1452–9. doi: 10.1111/j.1420-9101.2008.01611.x.

Fletcher, N. *et al.* (2021) 'A pheromone bouquet controls the reproductive behaviour of the male shore crab, *Carcinus maenas*', *Aquatic Ecology*. doi: 10.1007/s10452-021-09930-w.

Forsberg, J. and Wiklund, C. (1989) 'Mating in the afternoon: Time-saving in courtship and remating by females of a polyandrous butterfly *Pieris napi* L.', *Behavioral Ecology and Sociobiology*. Springer-Verlag, 25(5), pp. 349–356. doi: 10.1007/BF00302992.

Foster, S. P. *et al.* (1997) 'A Genetic Study on Pheromonal Communication in Two *Ctenopseustis* Moths', in *Insect Pheromone Research*. Springer US, pp. 514–524. doi: 10.1007/978-1-4615-6371-6_44.

Fraenkel, G. and Blewett, M. (1945) 'Linoleic acid, α-tocopherol and other fat-soluble substances as nutritional factors for insects', *Nature*. Nature Publishing Group, pp. 392–393. doi: 10.1038/155392c0.

Freeman, P. W. (2000) *Macroevolution in Microchiroptera: Recoupling morphology and ecology with phylogeny*, *digitalcommons.unl.edu*.

Freitas, A. V. L. and Brown, K. S. (2004) 'Phylogeny of the Nymphalidae (Lepidoptera)', *Systematic Biology*. Edited by T. Schultz. Oxford Academic, 53(3), pp. 363–383. doi: 10.1080/10635150490445670.

Frenkel, J., Vyverman, W. and Pohnert, G. (2014) 'Pheromone signaling during sexual reproduction in algae', *The Plant Journal*. Blackwell Publishing Ltd, 79(4), pp. 632–644. doi: 10.1111/tpj.12496.

Fuchs, M. S., Craig, G. B. and Despommier, D. D. (1969) 'The protein nature of the substance inducing female monogamy in *Aedes aegypti*', *Journal of Insect Physiology*. Pergamon, 15(4), pp. 701–709. doi: 10.1016/0022-1910(69)90265-0.

Fuchs, M. S., Craig, G. B. and Hiss, E. A. (1968) 'The biochemical basis of female monogamy in mosquitoes. I. Extraction of the active principle from *Aedes aegypti*', *Life Sciences*. Life Sci, 7(16 PART 2), pp. 835–839. doi: 10.1016/0024-3205(68)90114-8.

Fulda, M. *et al.* (2002) 'Two long-chain acyl-CoA synthetases from *Arabidopsis thaliana* involved in peroxisomal fatty acid β-oxidation', *Plant Journal*, 32(1), pp. 93–103. doi: 10.1046/j.1365-313X.2002.01405.x.

Gassmann, G., Müller, D. G. and Fritz, P. (1995) 'Pheromones in marine algae: A technical approach', *Helgoländer Meeresuntersuchungen*. Springer-Verlag, 49(1–4), pp. 797–804. doi: 10.1007/BF02368402.

Genner, M. J. *et al.* (2007) 'Reproductive isolation among deep-water cichlid fishes of Lake Malawi differing in monochromatic male breeding dress', *Molecular Ecology*. Mol Ecol, 16(3), pp. 651–662. doi: 10.1111/j.1365-294X.2006.03173.x.

Giglio, A. *et al.* (2009) 'The Defensive Secretion of *Carabus lefebvrei* Dejean 1826 Pupa (Coleoptera, Carabidae): Gland Ultrastructure and Chemical Identification', *Microscopy Research and Technique*. John Wiley & Sons, Ltd, 72(5), pp. 351–361. doi: 10.1002/jemt.20660.

Gilbert, L. E. (1972) 'Pollen Feeding and Reproductive Biology of *Heliconius* Butterflies', *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences, 69(6), pp. 1403–1407. doi: 10.1073/pnas.69.6.1403.

Gilbert, L. E. (1975) 'Ecological consequences of a coevolved mutualism between butterflies and plants.', in Gilbert, L. E. and Raven, P. R. (eds) *Coevolution of Animals and Plants*. Austin, Texas: University of Texas Press, pp. 210–240.

Gilbert, L. E. (1976) 'Postmating female odor in *Heliconius* butterflies: A male-contributed antiaphrodisiac?', *Science*. American Association for the Advancement of Science, 193(4251), pp. 419–420. doi: 10.1126/science.935877.

Ginzel, M. D. (2009) 'Olfactory Signals', in *Encyclopedia of Animal Behavior*. Elsevier Inc., pp. 584–588. doi: 10.1016/B978-0-08-045337-8.00268-0.

Gleason, J. M. and Ritchie, M. G. (1998) 'Evolution of Courtship Song and Reproductive Isolation in the *Drosophila willistoni* Species Complex: Do Sexual Signals Diverge the Most Quickly?', *Evolution*. John Wiley & Sons, Ltd (10.1111), 52(5), p. 1493. doi: 10.2307/2411319.

Glover, T. J., Tang, X. H. and Roelofs, W. L. (1987) 'Sex pheromone blend discrimination by male moths from E and Z strains of European corn borer', *Journal of Chemical Ecology*. Kluwer Academic Publishers-Plenum Publishers, 13(1), pp. 143–151. doi: 10.1007/BF01020358.

Gomendio, M., Harcourt, A. H. and Roldán, E. R. S. (1998) 'Sperm Competition in Mammals', in *Sperm Competition and Sexual Selection*. Elsevier, pp. 667–755. doi: 10.1016/b978-012100543-6/50041-6.

González-Rojas, M. F. *et al.* (2020) 'Chemical signals act as the main reproductive barrier between sister and mimetic *Heliconius* butterflies', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society Publishing, 287(1926), p. 20200587. doi: 10.1098/rspb.2020.0587.

González, M. A. *et al.* (2017) 'A temporal comparison of sex-aggregation pheromone gland content and dynamics of release in three members of the *Lutzomyia longipalpis* (Diptera: Psychodidae) species complex', *PLOS Neglected Tropical Diseases*. Edited by P. L. Oliveira. Public Library of Science, 11(12), p. e0006071. doi: 10.1371/journal.pntd.0006071.

Gosling, L. M. and Roberts, S. C. (2001) 'Scent-marking by male mammals: Cheat-proof signals to competitors and mates', *Advances in the Study of Behavior*. Academic Press Inc., 30, pp. 169–217. doi: 10.1016/s0065-3454(01)80007-3.

Gould, F. *et al.* (2010) 'Sexual isolation of male moths explained by a single pheromone response QTL containing four receptor genes', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 107(19), pp. 8660–8665. doi: 10.1073/pnas.0910945107.

Grammer, K., Fink, B. and Neave, N. (2005) 'Human pheromones and sexual attraction', *European Journal of Obstetrics and Gynecology and Reproductive Biology*. Elsevier Ireland Ltd, pp. 135–142. doi: 10.1016/j.ejogrb.2004.08.010.

Gray, D. A. and Cade, W. H. (2000) 'Sexual selection and speciation in field crickets', *Proceedings of the National Academy of Sciences of the United States of America*, 97(26), pp. 14449–14454. doi: 10.1073/pnas.97.26.14449.

Groot, A. T. *et al.* (2014) 'Within-population variability in a moth sex pheromone blend: Genetic basis and behavioural consequences', *Proceedings of the Royal Society B: Biological Sciences*. The Royal Society, 281(1779), pp. 20133054–20133054. doi: 10.1098/rspb.2013.3054.

Groot, A. T. *et al.* (2019) 'Within-population variability in a moth sex pheromone blend, part 2: Selection towards fixation', *Royal Society Open Science*. The Royal Society, 6(3), p. 182050. doi: 10.1098/rsos.182050.

Groot, A. T., Dekker, T. and Heckel, D. G. (2016) 'The Genetic Basis of Pheromone Evolution in

Moths', *Annual Review of Entomology*. Annual Reviews Inc., pp. 99–117. doi: 10.1146/annurev-ento-010715-023638.

Grula, J. W., McChesney, J. D. and Taylor, O. R. (1980) 'Aphrodisiac pheromones of the sulfur butterflies *Colias eurytheme* and *C. philodice* (Lepidoptera, Pieridae)', *Journal of Chemical Ecology*. Kluwer Academic Publishers-Plenum Publishers, 6(1), pp. 241–256. doi: 10.1007/BF00987543.

Von Hagen, K. B. and Kadereit, J. W. (2003) 'The diversification of *Halenia* (Gentianaceae): Ecological opportunity versus key innovation', *Evolution*. Society for the Study of Evolution, 57(11), pp. 2507–2518. doi: 10.1111/j.0014-3820.2003.tb01495.x.

Hagström, Å. K. *et al.* (2012) 'Semi-selective fatty acyl reductases from four Heliothine moths influence the specific pheromone composition', *PLoS ONE*. Edited by M. Renou. Public Library of Science, 7(5), p. e37230. doi: 10.1371/journal.pone.0037230.

Haley, C. S. and Knott, S. A. (1992) 'A simple regression method for mapping quantitative trait loci in line crosses using flanking markers', *Heredity*. Nature Publishing Group, 69(4), pp. 315–324. doi: 10.1038/hdy.1992.131.

Hansson, B. S. (1995) 'Olfaction in Lepidoptera', *Experientia*. Birkhäuser-Verlag, pp. 1003–1027. doi: 10.1007/BF01946910.

Happ, G. M. (1969) 'Multiple sex pheromones of the mealworm beetle, *Tenebrio molitor* L.', *Nature*. Nature Publishing Group, pp. 180–181. doi: 10.1038/222180a0.

Hardege, J. D. *et al.* (1998) 'Timing of reproduction in marine polychaetes: the role of sex pheromones', *Ecoscience*. Universite Laval, 5(3), pp. 395–404. doi: 10.1080/11956860.1998.11682477.

Harmon, L. J. *et al.* (2008) 'GEIGER: investigating evolutionary radiations', *Bioinformatics*. Narnia, 24(1), pp. 129–131. doi: 10.1093/bioinformatics/btm538.

Harmon, L. J. (2019) *Phylogenetic Comparative Methods*. doi: doi:10.32942/osf.io/e3xnr.

Hatfield, T. and Schluter, D. (1999) 'Ecological Speciation in Sticklebacks: Environment-Dependent Hybrid Fitness', *Evolution*. JSTOR, 53(3), p. 866. doi: 10.2307/2640726.

Hayashi, N. *et al.* (1985) 'The Scent substances of pierid butterflies (*Hebomoia glaucippe* Linnaeus) and the volatile components of their food plants (*Crataeva religiosa* Forst.)', *Zeitschrift fur Naturforschung - Section C Journal of Biosciences*, 40(1–2), pp. 47–50. doi: 10.1515/znc-1985-1-211.

He, P. *et al.* (2017) 'A reference gene set for sex pheromone biosynthesis and degradation genes from the diamondback moth, *Plutella xylostella*, based on genome and transcriptome digital gene expression analyses', *BMC Genomics*. BioMed Central Ltd., 18(1), pp. 1–19. doi: 10.1186/s12864-017-3592-y.

Heard, S. B. and Hauser, D. L. (1995) 'Key evolutionary innovations and their ecological mechanisms', *Historical Biology*. Taylor & Francis Group, 10(2), pp. 151–173. doi: 10.1080/10292389509380518.

Henneken, J. *et al.* (2015) 'Diet influences female signal reliability for male mate choice', *Animal Behaviour*. Academic Press, 108, pp. 215–221. doi: 10.1016/j.anbehav.2015.07.023.

Henneken, J. *et al.* (2017) 'Diet-mediated pheromones and signature mixtures can enforce signal reliability', *Frontiers in Ecology and Evolution*. Frontiers Media S. A, p. 145. doi: 10.3389/fevo.2016.00145.

Heuskin, S. *et al.* (2014) 'The composition of cuticular compounds indicates body parts, sex and age in the model butterfly *Bicyclus anynana* (Lepidoptera)', *Frontiers in Ecology and Evolution*. Frontiers

Media S. A, 2(JUL). doi: 10.3389/fevo.2014.00037.

Hildebrand, J. G. and Shepherd, G. M. (1997) 'Mechanisms of olfactory discrimination: Converging evidence for common principles across phyla', *Annual Review of Neuroscience*. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, pp. 595–631. doi: 10.1146/annurev.neuro.20.1.595.

Hill, A. S. *et al.* (1977) 'Sex pheromone components of the variegated leafroller moth, *Platynota flavedana*', *Journal of Chemical Ecology*. Kluwer Academic Publishers-Plenum Publishers, 3(4), pp. 369–376. doi: 10.1007/BF00988180.

Hoffman, J. I. *et al.* (2014) 'High-throughput sequencing reveals inbreeding depression in a natural population', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 111(10), pp. 3775–3780. doi: 10.1073/pnas.1318945111.

Honda, K. (1981) 'Larval osmeterial secretions of the swallowtails (*Papilio*)', *Journal of Chemical Ecology*. Kluwer Academic Publishers-Plenum Publishers, 7(6), pp. 1089–1113. doi: 10.1007/BF00987631.

Hosken, D. J., Archer, C. R. and Mank, J. E. (2019) 'Sexual conflict', *Current Biology*. Cell Press, pp. R451–R455. doi: 10.1016/j.cub.2019.03.052.

Howard, R. W. and Blomquist, G. J. (1982) 'Chemical Ecology and Biochemistry of Insect Hydrocarbons', *Annual Review of Entomology*, 27(1), pp. 149–172. doi: 10.1146/annurev.en.27.010182.001053.

Howard, R. W. and Blomquist, G. J. (2005) 'Ecological, behavioral, and biochemical aspects of insect hydrocarbons', *Annual Review of Entomology*, pp. 371–393. doi: 10.1146/annurev.ento.50.071803.130359.

Hübner, G. *et al.* (2002) 'Mandibular gland secretions in alloxystine wasps (Hymenoptera, Cynipoidea, Charipidae): Do ecological or phylogenetical constraints influence occurrence or composition?', *Biochemical Systematics and Ecology*. Pergamon, 30(6), pp. 505–523. doi: 10.1016/S0305-1978(01)00137-5.

Hunter, J. P. (1998) 'Key innovations and the ecology of macroevolution', *Trends in Ecology and Evolution*. Elsevier Current Trends, 13(1), pp. 31–36. doi: 10.1016/S0169-5347(97)01273-1.

Hunter, J. P. and Jernvall, J. (1995) 'The hypocone as a key innovation in mammalian evolution', *Proceedings of the National Academy of Sciences of the United States of America*, 92(23), pp. 10718–10722. doi: 10.1073/pnas.92.23.10718.

Hurst, J. L. *et al.* (2001) 'Individual recognition in mice mediated by major urinary proteins', *Nature*, 414(6864), pp. 631–634. doi: 10.1038/414631a.

Jacob, E. Ben *et al.* (2004) 'Bacterial linguistic communication and social intelligence', *Trends in Microbiology*. Elsevier Ltd, 12(8), pp. 366–372. doi: 10.1016/j.tim.2004.06.006.

Jaenicke, L. and Starr, R. C. (1996) 'The lurlenes, a new class of plastoquinone-related mating pheromones from *Chlamydomonas allensworthii* (Chlorophyceae)', *European Journal of Biochemistry*, 241(2), pp. 581–585. doi: 10.1111/j.1432-1033.1996.00581.x.

Jennions, M. D. and Petrie, M. (2000) 'Why do females mate multiply? A review of the genetic benefits', *Biological Reviews of the Cambridge Philosophical Society*. Wiley-Blackwell, 75(1), pp. 21–64. doi: 10.1017/s0006323199005423.

Jiggins, C. D. *et al.* (2001) 'Reproductive isolation caused by colour pattern mimicry.', *Nature*, 411(6835), pp. 302–5. doi: 10.1038/35077075.

Jiggins, C. D. (2008) 'Ecological Speciation in Mimetic Butterflies', *BioScience*. Oxford University Press, 58(6), pp. 541–548. doi: 10.1641/B580610.

Jiggins, C. D. *et al.* (2008) 'Review. Hybrid trait speciation and *Heliconius* butterflies', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, pp. 3047–3054. doi: 10.1098/rstb.2008.0065.

Jiggins, C. D. and Lamas, G. (2017) *The Ecology and Evolution of* Heliconius *Butterflies*, *The Ecology and Evolution of Heliconius Butterflies*. Oxford University Press. doi: 10.1093/acprof:oso/9780199566570.001.0001.

Johansson, B. G. and Jones, T. M. (2007) 'The role of chemical communication in mate choice', *Biological Reviews*. John Wiley & Sons, Ltd, 82(2), pp. 265–289. doi: 10.1111/j.1469-185X.2007.00009.x.

Jones, A. G., Adams, E. M. and Arnold, S. J. (2002) 'Topping off: A mechanism of first-male sperm precedence in a vertebrate', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 99(4), pp. 2078–2081. doi: 10.1073/pnas.042510199.

Jones, T. C. and Parker, P. G. (2008) 'First male sperm precedence in multiply-mated females of the cooperative spider *Anelosimus studiosus* (Araneae, Theridiidae)', *Journal of Arachnology*. American Museum of Natural History, 36(3), pp. 527–532. doi: 10.1636/St06-63.1.

Joron, M. *et al.* (1999) 'Variable Selection and the Coexistence of Multiple mimetic forms of the Butterfly *Heliconius numata*', *Evolutionary Ecology*. Kluwer Academic Publishers, 13(7–8), pp. 721–754. doi: 10.1023/A:1010875213123.

Jousselin, E., Van Noort, S. and Greeff, J. M. (2004) 'Labile male morphology and intraspecific male polymorphism in the *Philotrypesis* fig wasps', *Molecular Phylogenetics and Evolution*, 33(3), pp. 706–718. doi: 10.1016/j.ympev.2004.08.008.

Ju, Q. *et al.* (2017) 'Plant Volatiles Increase Sex Pheromone Attraction of *Holotrichia parallela* (Coleoptera: Scarabaeoidea)', *Journal of Chemical Ecology*. Springer New York LLC, 43(3), pp. 236–242. doi: 10.1007/s10886-017-0823-2.

Jurenka, R. (2004) 'Insect Pheromone Biosynthesis', in, pp. 97–132. doi: 10.1007/b95450.

Jurenka, R. *et al.* (2017) 'Biochemistry and Molecular Biology of Pheromone Production', in *Reference Module in Life Sciences*, pp. 1–48. doi: 10.1016/b978-0-12-809633-8.04037-1.

Jurenka, R. A. *et al.* (2003) 'Sex pheromone biosynthetic pathway for disparlure in the gypsy moth, *Lymantria dispar*', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 100(3), pp. 809–814. doi: 10.1073/pnas.0236060100.

Jurenka, R. A. and Roelofs, W. L. (1993) 'Biosynthesis and endocrine regulation of fatty acid derived sex pheromones in moths', in Stanley, D. W. and Nelson, D. R. (eds) *Insect Lipids: Chemistry, Biochemistry and Biology*. Lincoln, Nebraska: University of Nebraska Press, pp. 353–380.

Karl, I., Heuskin, S. and Fischer, K. (2013) 'Dissecting the mechanisms underlying old male mating advantage in a butterfly', *Behavioral Ecology and Sociobiology*. Springer Verlag, 67(5), pp. 837–849. doi: 10.1007/s00265-013-1507-8.

Karlson, P. and Lüscher, M. (1959) '"Pheromones": A new term for a class of biologically active substances', *Nature*. Nature Publishing Group, 183(4653), pp. 55–56. doi: 10.1038/183055a0.

Karlsson, B. (1994) 'Feeding Habits and Change of Body Composition with Age in Three Nymphalid Butterfly Species', *Oikos*. JSTOR, 69(2), p. 224. doi: 10.2307/3546142.

Kawecki, T. J. (no date) 'Genetical theories of sympatric speciation.', in Dieckmann, U. et al. (eds) *Adaptive Speciation.* Cambridge: Cambridge University Press, pp. 36–53.

Keegans, S. J. *et al.* (1993) 'Volatile glandular secretions of three species of new world army ants, *Eciton burchelli, Labidus coecus*, and *Labidus praedator*', *Journal of Chemical Ecology*. Kluwer Academic Publishers-Plenum Publishers, 19(11), pp. 2705–2719. doi: 10.1007/BF00980702.

Klein, A. L. and de Araújo, A. M. (2010) 'Courtship behavior of *Heliconius erato phyllis* (Lepidoptera, Nymphalidae) towards virgin and mated females: Conflict between attraction and repulsion signals?', *Journal of Ethology*, 28(3), pp. 409–420. doi: 10.1007/s10164-010-0209-1.

Klun, J. A. *et al.* (1973) 'Insect sex pheromones: Minor amount of opposite geometrical isomer critical to attraction', *Science*. American Association for the Advancement of Science , 181(4100), pp. 661–663. doi: 10.1126/science.181.4100.661.

Knudsen, J. T. *et al.* (2006) 'Diversity and distribution of floral scent', *Botanical Review*, pp. 1–120. doi: 10.1663/0006-8101(2006)72[1:DADOFS]2.0.CO;2.

Kochansky, J. *et al.* (1975) 'Sex pheromone of the european corn borer, *Ostrinia nubilalis* (Lepidoptera: Pyralidae), in New York', *Journal of Chemical Ecology*. Kluwer Academic Publishers-Plenum Publishers, 1(2), pp. 225–231. doi: 10.1007/BF00987871.

Konow, N. *et al.* (2017) 'Decoupled diversification dynamics of feeding morphology following a major functional innovation in marine butterflyfishes', *Proceedings of the Royal Society B: Biological Sciences*. The Royal Society, 284(1860), p. 20170906. doi: 10.1098/rspb.2017.0906.

Kopena, R. *et al.* (2011) 'Vitamin E Supplementation Increases the Attractiveness of Males' Scent for Female European Green Lizards', *PLoS ONE*. Edited by S. R. Proulx. Public Library of Science, 6(4), p. e19410. doi: 10.1371/journal.pone.0019410.

Kozak, K. M. *et al.* (2015) 'Multilocus species trees show the recent adaptive radiation of the mimetic *Heliconius* butterflies', *Systematic Biology*. Oxford University Press, 64(3), pp. 505–524. doi: 10.1093/sysbio/syv007.

Kronforst, M. R. *et al.* (2006) 'Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless', *Proceedings of the National Academy of Sciences of the United States of America*, 103(17), pp. 6575–6580. doi: 10.1073/pnas.0509685103.

Kryvokhyzha, D. (2014) 'Whole genome resequencing of *Heliconius* butterflies revolutionizes our view of the level of admixture between species', *Master's Thesis in Evolutionary Biology, Uppsala University & Harvard University*, (January), p. 30.

Kukuk, P. (1985) 'Evidence for an antiaphrodisiac in the sweat bee *Lasioglossum (dialictus) zephyrum*', *Science*. American Association for the Advancement of Science, 227(4687), pp. 656–657. doi: 10.1126/science.3969557.

Laidre, M. E. and Johnstone, R. A. (2013) 'Animal signals', *Current Biology*. Cell Press, pp. R829–R833. doi: 10.1016/j.cub.2013.07.070.

Lankinen, Å. and Green, K. K. (2015) 'Using theories of sexual selection and sexual conflict to improve our understanding of plant ecology and evolution', *AoB PLANTS*. Oxford University Press. doi: 10.1093/aobpla/plv008.

Larsdotter-Mellström, H. *et al.* (2012) 'Timing of Male Sex Pheromone Biosynthesis in a Butterfly - Different Dynamics under Direct or Diapause Development', *Journal of Chemical Ecology*. Springer, 38(5), pp. 584–591. doi: 10.1007/s10886-012-0126-6.

Lassance, J. M. *et al.* (2010) 'Allelic variation in a fatty-acyl reductase gene causes divergence in moth

sex pheromones', *Nature*. Nature Research, 466(7305), pp. 486–489. doi: 10.1038/nature09058.

Laturney, M. and Billeter, J. C. (2016) '*Drosophila melanogaster* females restore their attractiveness after mating by removing male anti-aphrodisiac pheromones', *Nature Communications*. Nature Publishing Group, 7. doi: 10.1038/ncomms12322.

Laturney, M., van Eijk, R. and Billeter, J.-C. (2018) 'Last male sperm precedence is modulated by female remating rate in *Drosophila melanogaster*', *Evolution Letters*. Wiley, 2(3), pp. 180–189. doi: 10.1002/evl3.50.

Li, C. *et al.* (2017) 'Role of visual and olfactory cues in sex recognition in butterfly *Cethosia cyane cyane*', *Scientific Reports*. Nature Publishing Group, 7(1), p. 5033. doi: 10.1038/s41598-017-04721-6.

Liedo, P. *et al.* (2013) 'Effect of post-teneral diets on the performance of sterile *Anastrepha ludens* and *Anastrepha obliqua* fruit flies', *Journal of Applied Entomology*. John Wiley & Sons, Ltd, 137(SUPPL.1), pp. 49–60. doi: 10.1111/j.1439-0418.2010.01568.x.

Liénard, M. A. *et al.* (2008) 'Key biosynthetic gene subfamily recruited for pheromone production prior to the extensive radiation of Lepidoptera'. doi: 10.1186/1471-2148-8-270.

Liénard, M. A. *et al.* (2014) 'Sex pheromone biosynthetic pathways are conserved between moths and the butterfly *Bicyclus anynana*', *Nature Communications*, 5(3957). doi: 10.1038/ncomms4957.

Linstrom, P. J. and Mallard, W. G. (eds) (2021) *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*. National Institute of Standards and Technology, Gaithersburg MD, 20899. doi: https://doi.org/10.18434/T4D303.

Livraghi, L. *et al.* (2021) 'Cortex cis-regulatory switches establish scale colour identity and pattern diversity in *Heliconius*', *eLife*. eLife Sciences Publications Ltd, 10. doi: 10.7554/eLife.68549.

Löfqvist, J. (1976) 'Formic acid and saturated hydrocarbons as alarm pheromones for the ant *Formica rufa*', *Journal of Insect Physiology*. Pergamon, 22(10), pp. 1331–1346. doi: 10.1016/0022-1910(76)90155-4.

Löfstedt, C. *et al.* (1989) 'Diet related courtship success in the Oriental fruit moth, *Grapholitha molesta* (Tortricidae)', *Oikos*. WileyNordic Society Oikos, 55(3), pp. 402–408. doi: 10.2307/3565601.

Löfstedt, C., Herrebout, W. M. and Menken, S. B. J. (1991) 'Sex pheromones and their potential role in the evolution of reproductive isolation in small ermine moths (Yponomeutidae)', *Chemoecology*. Birkhäuser-Verlag, 2(1), pp. 20–28. doi: 10.1007/BF01240662.

Löfstedt, C. W., Wahlberg, N. and Millar, J. G. (2019) '4. Evolutionary Patterns of Pheromone Diversity in Lepidoptera', in *Pheromone Communication in Moths*. University of California Press, pp. 43–78. doi: 10.1525/9780520964433-005.

López, P., Aragón, P. and Martín, J. (2003) 'Responses of female lizards, *Lacerta monticola*, to males' chemical cues reflect their mating preference for older males', *Behavioral Ecology and Sociobiology*, 55(1), pp. 73–79. doi: 10.1007/s00265-003-0675-3.

Losos, J. B. (2008) 'Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species.', *Ecology letters*, 11(10), pp. 995–1003. doi: 10.1111/j.1461-0248.2008.01229.x.

Losos, J. B. (2010) 'Adaptive radiation, ecological opportunity, and evolutionary determinism : American society of naturalists E. O. Wilson award address', *American Naturalist*. The University of Chicago Press , 175(6), pp. 623–639. doi: 10.1086/652433.

Lovette, I. J., Bermingham, E. and Ricklefs, R. E. (2002) 'Clade-specific morphological diversification

and adaptive radiation in Hawaiian songbirds.', *Proceedings. Biological sciences*. The Royal Society, 269(1486), pp. 37–42. doi: 10.1098/rspb.2001.1789.

Lowry, D. B., Rockwood, R. C. and Willis, J. H. (2008) 'Ecological reproductive isolation of coast and inland races of *Mimulus guttatus*', *Evolution*, 62(9), pp. 2196–2214. doi: 10.1111/j.1558-5646.2008.00457.x.

Lundgren, L. and Bergström, G. (1975) 'Wing scents and scent-released phases in the courtship behavior of *Lycaeides argyrognomon* (Lepidoptera: Lycaenidae)', *Journal of Chemical Ecology*. Kluwer Academic Publishers-Plenum Publishers, 1(4), pp. 399–412. doi: 10.1007/BF00988581.

Luporini, P., Alimenti, C. and Vallesi, A. (2015) 'Ciliate pheromone structures and activity: a review', *Italian Journal of Zoology*. Taylor and Francis Ltd., 82(1), pp. 3–14. doi: 10.1080/11250003.2014.976282.

MacLean, M. *et al.* (2018) 'Mountain pine beetle (Dendroctonus ponderosae) CYP4Gs convert long and short chain alcohols and aldehydes to hydrocarbons', *Insect Biochemistry and Molecular Biology*. Elsevier Ltd, 102, pp. 11–20. doi: 10.1016/j.ibmb.2018.09.005.

Mahler, D. L. *et al.* (2010) 'Ecological opportunity and the rate of morphological evolution in the diversification of greater Antillean anoles', *Evolution*. John Wiley & Sons, Ltd, 64(9), pp. 2731–2745. doi: 10.1111/j.1558-5646.2010.01026.x.

Maia, R., Rubenstein, D. R. and Shawkey, M. D. (2013) 'Key ornamental innovations facilitate diversification in an avian radiation', *Proceedings of the National Academy of Sciences*, 110(26), pp. 10687–10692. doi: 10.1073/pnas.1220784110.

Mallard, W. G. and Reed, J. (1997) *Automated Mass Spectral Deconvolution & Identification System AMDIS-User Guide*.

Mallet, J. (1995) 'A species definition for the modern synthesis', *Trends in Ecology & Evolution*. Elsevier Current Trends, 10(7), pp. 294–299. doi: 10.1016/0169-5347(95)90031-4.

Mallet, J. and Dasmahapatra, K. K. (2012) 'Hybrid zones and the speciation continuum in *Heliconius* butterflies', *Molecular Ecology*. Mol Ecol, pp. 5643–5645. doi: 10.1111/mec.12058.

Mallet, J. and Joron, M. (1999) 'Evolution of diversity in warning color and mimicry: Polymorphisms, shifting balance, and speciation', *Annual Review of Ecology and Systematics*, 30, pp. 201–233. doi: 10.1146/annurev.ecolsys.30.1.201.

Mallet, J. L. B. and Gilbert, L. E. (1995) 'Why are there so many mimicry rings? Correlations between habitat, behaviour and mimicry in *Heliconius* butterflies', *Biological Journal of the Linnean Society*. John Wiley & Sons, Ltd, 55(2), pp. 159–180. doi: 10.1111/j.1095-8312.1995.tb01057.x.

Mann, F. *et al.* (2017) 'The Scent Chemistry of *Heliconius* Wing Androconia', *Journal of Chemical Ecology*. Springer US, 43(9), pp. 843–857. doi: 10.1007/s10886-017-0867-3.

Martin, A. *et al.* (2012) 'Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 109(31), pp. 12632–12637. doi: 10.1073/pnas.1204800109.

Martin, E. and Hine, R. (2008) *'Semiochemical'*, *A Dictionary of Biology*. Oxford University Press. doi: 10.1093/acref/9780199204625.001.0001.

Martin, S. H. *et al.* (2013) 'Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies', *Genome Research*. Cold Spring Harbor Laboratory Press, 23(11), pp. 1817–1828. doi: 10.1101/gr.159426.113.

Martin, S. H. and Van Belleghem, S. M. (2017) 'Exploring evolutionary relationships across the genome using topology weighting', *Genetics*. Genetics Society of America, 206(1), pp. 429–438. doi: 10.1534/genetics.116.194720.

Mason, N. A. *et al.* (2017) 'Song evolution, speciation, and vocal learning in passerine birds', *Evolution*. Society for the Study of Evolution, 71(3), pp. 786–796. doi: 10.1111/evo.13159.

Masta, S. E. and Maddison, W. P. (2002) 'Sexual selection driving diversification in jumping spiders', *Proceedings of the National Academy of Sciences of the United States of America*, 99(7), pp. 4442–4447. doi: 10.1073/pnas.072493099.

Matysioková, B. *et al.* (2017) 'The evolution of feather coloration and song in Old World orioles (genus Oriolus)', *Journal of Avian Biology*. Blackwell Publishing Ltd, 48(7), pp. 1015–1024. doi: 10.1111/jav.01175.

Mavárez, J. *et al.* (2006) 'Speciation by hybridization in *Heliconius* butterflies', *Nature*, 441(7095), pp. 868–871. doi: 10.1038/nature04738.

Mayr, E. (1942) *Systematics and the origin of species, from the viewpoint of a zoologist*. New York: Columbia University Press.

Mayr, E. (1963) *Animal Species and Evolution*, *Animal Species and Evolution*. HUP. doi: 10.4159/harvard.9780674865327.

Mazo-Vargas, A. *et al.* (2017) 'Macroevolutionary shifts of WntA function potentiate butterfly wing-pattern diversity', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 114(40), pp. 10701–10706. doi: 10.1073/pnas.1708149114.

McElfresh, J. S. and Millar, J. G. (2001) 'Geographic variation in the pheromone system of the saturniid moth *Hemileuca eglanterina*', *Ecology*, 82(12), pp. 3505–3518. doi: 10.2307/2680168.

McMillan, W. O. *et al.* (2020) 'From Patterning Genes to Process: Unraveling the Gene Regulatory Networks That Pattern *Heliconius* Wings', *Frontiers in Ecology and Evolution*. Frontiers Media S.A. doi: 10.3389/fevo.2020.00221.

Melo, M. C. *et al.* (2009) 'Assortative mating preferences among hybrids offers a route to hybrid speciation', *Evolution*. Society for the Study of Evolution, 63(6), pp. 1660–1665. doi: 10.1111/j.1558-5646.2009.00633.x.

Mendelson, T. C. and Shaw, K. L. (2005) 'Sexual behaviour: Rapid speciation in an anthropod', *Nature*. Nature Publishing Group, 433(7024), pp. 375–376. doi: 10.1038/433375a.

Mendoza-Cuenca, L. and Macías-Ordóñez, R. (2010) 'Female asynchrony may drive disruptive sexual selection on male mating phenotypes in a *Heliconius* butterfly', *Behavioral Ecology*. Oxford University Press, 21(1), pp. 144–152. doi: 10.1093/beheco/arp163.

Mérot, C. *et al.* (2013) 'Genetic differentiation without mimicry shift in a pair of hybridizing *Heliconius* species (Lepidoptera: Nymphalidae)', *Biological Journal of the Linnean Society*, 109(4), pp. 830–847. doi: 10.1111/bij.12091.

Mérot, C. *et al.* (2015) 'Beyond magic traits: Multimodal mating cues in *Heliconius* butterflies', *Evolution*, 69(11), pp. 2891–2904. doi: 10.1111/evo.12789.

Mérot, C. *et al.* (2017) 'What shapes the continuum of reproductive isolation? Lessons from *Heliconius* butterflies', *Proceedings of the Royal Society B: Biological Sciences*. The Royal Society, 284(1856), p. 20170335. doi: 10.1098/rspb.2017.0335.

Merrill, R. M. *et al.* (2011) 'Pervasive genetic associations between traits causing reproductive

isolation in *Heliconius* butterflies', in *Proceedings of the Royal Society B: Biological Sciences*. The Royal Society, pp. 511–518. doi: 10.1098/rspb.2010.1493.

Merrill, R. M. *et al.* (2012) 'Disruptive ecological selection on a mating cue.', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 279(1749), pp. 4907–4913. doi: 10.1098/rspb.2012.1968.

Merrill, R. M. *et al.* (2015) 'The diversification of Heliconius butterflies: What have we learned in 150 years?', *Journal of Evolutionary Biology*. J Evol Biol, 28(8), pp. 1417–1438. doi: 10.1111/jeb.12672.

Merrill, R. M., Chia, A. and Nadeau, N. J. (2014) 'Divergent warning patterns contribute to assortative mating between incipient *Heliconius* species', *Ecology and Evolution*. John Wiley and Sons Ltd, 4(7), pp. 911–917. doi: 10.1002/ece3.996.

Merwin, J. T., Seeholzer, G. F. and Smith, B. T. (2020) 'Macroevolutionary bursts and constraints generate a rainbow in a clade of tropical birds', *BMC Evolutionary Biology*. BioMed Central Ltd., 20(1). doi: 10.1186/s12862-020-1577-y.

Michel, A. P. *et al.* (2010) 'Widespread genomic divergence during sympatric speciation', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 107(21), pp. 9724–9729. doi: 10.1073/pnas.1000939107.

Miller, A. (1949) 'Some ecologic and morphological considerations in the evolution of higher taxonomic categories.', in Mayr, E. and Schulz, E. (eds) *Ornithologie als biologische wissenschaft*. Heidelberg, Germany: Carl Winter, pp. 84–88.

Mishler, B. and Theriot, E. (2000) 'The Phylogenetic Species Concept (sensu Mishler and Theriot): Monophyly, Apomorphy, and Phylogenetic Species Concepts.', in Wheeler, Q. and Rudolf, M. (eds) *Species Concepts and Phylogenetic Theory: A Debate*. New York: Columbia University Press, pp. 44–54.

Moadeli, T. *et al.* (2020) 'Effects of fatty acids and vitamin E in larval diets on development and performance of Queensland fruit fly', *Journal of Insect Physiology*. Elsevier Ltd, 125, p. 104058. doi: 10.1016/j.jinsphys.2020.104058.

Montgomery, S. H., Merrill, R. M. and Ott, S. R. (2016) 'Brain composition in *Heliconius* butterflies, posteclosion growth and experience-dependent neuropil plasticity', *Journal of Comparative Neurology*. Wiley-Liss Inc., 524(9), pp. 1747–1769. doi: 10.1002/cne.23993.

Monti, L. *et al.* (1997) 'A genetic analysis of some components of reproductive isolation between two closely related species, *Spodoptera latifascia* (Walker) and *S. descoinsi* (Lalanne-Cassou and Silvain) (Lepidoptera: Noctuidae)', *Journal of Evolutionary Biology*. Wiley, 10(1), pp. 121–134. doi: 10.1046/j.1420-9101.1997.10010121.x.

Moritz, R. F. A. and Bürgin, H. (2010) 'Group Response to Alarm Pheromones in Social Wasps and the Honeybee', *Ethology*. John Wiley & Sons, Ltd, 76(1), pp. 15–26. doi: 10.1111/j.1439-0310.1987.tb00668.x.

Moser, J. C., Brownlee, R. C. and Silverstein, R. (1968) 'Alarm pheromones of the ant *Atta texana*', *Journal of Insect Physiology*. Pergamon, 14(4), pp. 529–535. doi: 10.1016/0022-1910(68)90068-1.

Moto, K. *et al.* (2003) 'Pheromone gland-specific fatty-acyl reductase of the silkmoth, *Bombyx mori*', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 100(16), pp. 9156–9161. doi: 10.1073/pnas.1531993100.

Moto, K. *et al.* (2004) 'Involvement of a bifunctional fatty-acyl desaturase in the biosynthesis of the silkmoth, *Bombyx mori*, sex pheromone', *Proceedings of the National Academy of Sciences of the*

United States of America. National Academy of Sciences, 101(23), pp. 8631–8636. doi: 10.1073/pnas.0402056101.

Mozuraitis, R. *et al.* (2019) 'Anti-aphrodisiac pheromone, a renewable signal in adult butterflies', *Scientific Reports*. Nature Publishing Group, 9(1), pp. 1–12. doi: 10.1038/s41598-019-50838-1.

Myers, J. (1972) 'Pheromones and Courtship Behavior in Butterflies', *American Zoologist*. Oxford Academic, 12(3), pp. 545–551. doi: 10.1093/icb/12.3.545.

Nadeau, N. J. *et al.* (2012) 'Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, 367(1587), pp. 343–353. doi: 10.1098/rstb.2011.0198.

Nadeau, N. J. *et al.* (2016) 'The gene cortex controls mimicry and crypsis in butterflies and moths', *Nature*, 534. doi: 10.1038/nature17961.

Naisbit, R. E., Jiggins, C. D. and Mallet, J. (2001) 'Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *Heliconius melpomene*', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 268(1478), pp. 1849–1854. doi: 10.1098/rspb.2001.1753.

Ng, T. P. T. *et al.* (2013) 'Snails and their trails: The multiple functions of trail-following in gastropods', *Biological Reviews*, 88(3), pp. 683–700. doi: 10.1111/brv.12023.

Nieberding, C. M. *et al.* (2008) 'The Male Sex Pheromone of the Butterfly *Bicyclus anynana*: Towards an Evolutionary Analysis', *PLoS ONE*. Edited by M. Somers. Public Library of Science, 3(7), p. e2751. doi: 10.1371/journal.pone.0002751.

Nieberding, C. M. *et al.* (2012) 'Cracking the olfactory code of a butterfly: the scent of ageing', *Ecology Letters*. John Wiley & Sons, Ltd, 15(5), pp. 415–424. doi: 10.1111/j.1461-0248.2012.01748.x.

Nieuwenhuis, B. P. S. and Aanen, D. K. (2012) 'Sexual selection in fungi', *Journal of Evolutionary Biology*. John Wiley & Sons, Ltd, pp. 2397–2411. doi: 10.1111/jeb.12017.

Nishida, R. *et al.* (1996) 'Male sex pheromone of a giant danaine butterfly, *Idea leuconoe*', *Journal of Chemical Ecology*, 22(5), pp. 949–972. doi: 10.1007/BF02029947.

Noor, M. A. F. (1999) 'Reinforcement and other consequences of sympatry', *Heredity*. John Wiley & Sons, Ltd (10.1111), 83(5), pp. 503–508. doi: 10.1046/J.1365-2540.1999.00632.X.

Norman, V. C. *et al.* (2017) 'Alarm Pheromone Composition and Behavioral Activity in Fungus-Growing Ants', *Journal of Chemical Ecology*. Springer New York LLC, 43(3), pp. 225–235. doi: 10.1007/s10886-017-0821-4.

Nosil, P. and Feder, J. L. (2012) 'Widespread yet heterogeneous genomic divergence', *Molecular Ecology*. John Wiley & Sons, Ltd, 21(12), pp. 2829–2832. doi: 10.1111/j.1365-294X.2012.05580.x.

Nosil, P., Funk, D. J. and Ortiz-Barrientos, D. (2009) 'Divergent selection and heterogeneous genomic divergence', *Molecular Ecology*. John Wiley & Sons, Ltd, pp. 375–402. doi: 10.1111/j.1365-294X.2008.03946.x.

Nosil, P. and Schluter, D. (2011) 'The genes underlying the process of speciation', *Trends in Ecology & Evolution*, 26, pp. 160–167. doi: 10.1016/j.tree.2011.01.001.

O'Callaghan, S. *et al.* (2012) 'PyMS: a Python toolkit for processing of gas chromatography-mass spectrometry (GC-MS) data. Application and comparative study of selected tools', *BMC Bioinformatics*. BioMed Central, 13(1), p. 115. doi: 10.1186/1471-2105-13-115.

Ohnishi, A. *et al.* (2011) 'Hormone signaling linked to silkmoth sex pheromone biosynthesis involves Ca2+/calmodulin-dependent protein kinase II-mediated phosphorylation of the insect PAT family protein Bombyx mori lipid storage droplet protein-1 (BmLsd1)', *Journal of Biological Chemistry*, 286(27), pp. 24101–24112. doi: 10.1074/jbc.M111.250555.

Oksanen, J., Guillaume Blanchet, F. and Kindt, R. (2019) *CRAN - Package vegan*. Available at: http://cran.r-project.org/web/packages/vegan/index.html (Accessed: 17 June 2020).

Ômura, H. and Honda, K. (2011) 'Pungent odor of the adult skipper butterfly *Erynnis montanus* (Lepidoptera: Hesperiidae)', *Applied Entomology and Zoology*. Springer, 46(2), pp. 281–286. doi: 10.1007/s13355-011-0039-9.

Ord, T. J. and Stamps, J. A. (2009) 'Species Identity Cues in Animal Communication', *American Naturalist*, 174(4), pp. 585–593. doi: 10.1086/605372.

Orr, H. A. (1995) 'The population genetics of speciation: The evolution of hybrid incompatibilities', *Genetics*. Oxford University Press, 139(4), pp. 1805–1813. doi: 10.1093/genetics/139.4.1805.

Ortiz-Barrientos, D., Grealy, A. and Nosil, P. (2009) 'The Genetics and Ecology of Reinforcement', *Annals of the New York Academy of Sciences*. Blackwell Publishing Inc., 1168(1), pp. 156–182. doi: 10.1111/j.1749-6632.2009.04919.x.

Painter, S. D. *et al.* (2004) 'Structural and functional analysis of *Aplysia* attractins, a family of water-borne protein pheromones with interspecific attractiveness', *Proceedings of the National Academy of Sciences of the United States of America*, 101(18), pp. 6929–6933. doi: 10.1073/pnas.0306339101.

Palacio Cortés, A. M. *et al.* (2010) 'Geographic variation of sex pheromone and mitochondrial DNA in *Diatraea saccharalis* (Fab., 1794) (Lepidoptera: Crambidae)', *Journal of Insect Physiology*. Pergamon, 56(11), pp. 1624–1630. doi: 10.1016/j.jinsphys.2010.06.005.

Panhuis, T. M. *et al.* (2001) 'Sexual selection and speciation', *Trends in Ecology and Evolution*. Elsevier, pp. 364–371. doi: 10.1016/S0169-5347(01)02160-7.

Pardo-Diaz, C. *et al.* (2012) 'Adaptive introgression across species boundaries in *Heliconius* butterflies', *PLoS Genetics*, 8(6). doi: 10.1371/journal.pgen.1002752.

Parker, G. A. (2006) 'Sexual conflict over mating and fertilization: An overview', in *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, pp. 235–259. doi: 10.1098/rstb.2005.1785.

Parker, M. R. and Mason, R. T. (2009) 'Low temperature dormancy affects the quantity and quality of the female sexual attractiveness pheromone in red-sided garter snakes', *Journal of Chemical Ecology*, 35(10), pp. 1234–1241. doi: 10.1007/s10886-009-9699-0.

Pearse, W. D. *et al.* (2018) 'Global macroevolution and macroecology of passerine song', *Evolution*. Society for the Study of Evolution, 72(4), pp. 944–960. doi: 10.1111/evo.13450.

Peña, A. *et al.* (1988) 'Sex pheromone of European corn borer. - *Ostrinia nubilalis*: Polymorphism in various laboratory and field strains', *Journal of Chemical Ecology*, 14(5), pp. 1359–1366. doi: 10.1007/BF01020140.

Pfennig, K. S. and Pfennig, D. W. (2009) 'Character displacement: ecological and reproductive responses to a common evolutionary problem', *Quarterly Review of Biology*, pp. 253–276. doi: 10.1086/605079.

Pfennig, K. S. and Rice, A. M. (2014) 'Reinforcement generates reproductive isolation between neighbouring conspecific populations of spadefoot toads', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society of London, 281(1789), p. 20140949. doi: 10.1098/rspb.2014.0949.

Pinheiro de Castro, É. C. *et al.* (2019) 'Sequestration and biosynthesis of cyanogenic glucosides in passion vine butterflies and consequences for the diversification of their host plants', *Ecology and Evolution*. John Wiley and Sons Ltd, 9(9), pp. 5079–5093. doi: 10.1002/ece3.5062.

Podos, J. (2001) 'Correlated evolution of morphology and vocal signal structure in Darwin's finches', *Nature*. Nature Publishing Group, 409(6817), pp. 185–188. doi: 10.1038/35051570.

Polechová, J. and Storch, D. (2019) 'Ecological Niche', in *Encyclopedia of Ecology*. Elsevier, pp. 72–80. doi: 10.1016/b978-0-12-409548-9.11113-3.

Polidori, C. *et al.* (2017) 'Post-mating shift towards longer-chain cuticular hydrocarbons drastically reduces female attractiveness to males in a digger wasp', *Journal of Insect Physiology*. Elsevier Ltd, 100, pp. 119–127. doi: 10.1016/j.jinsphys.2017.05.001.

Pontes, G. B. *et al.* (2008) 'Metasternal gland volatiles and sexual communication in the triatomine bug, *Rhodnius prolixus*', *Journal of Chemical Ecology*. Springer, 34(4), pp. 450–457. doi: 10.1007/s10886-008-9431-5.

Price, C. S. C., Dyer, K. A. and Coyne, J. A. (1999) 'Sperm competition between *Drosophila* males involves both displacement and incapacitation', *Nature*. Nature Publishing Group, 400(6743), pp. 449–452. doi: 10.1038/22755.

Puttick, M. N. *et al.* (2020) 'MOTMOT: Models of trait macroevolution on trees (an update)', *Methods in Ecology and Evolution*. British Ecological Society, 11(3), pp. 464–471. doi: 10.1111/2041-210X.13343.

Quattrocchio, F. *et al.* (1999) 'Molecular Analysis of the *anthocyanin2* Gene of Petunia and Its Role in the Evolution of Flower Color', *The Plant Cell*. American Society of Plant Biologists, 11(8), pp. 1433–1444. doi: 10.1105/tpc.11.8.1433.

Queste, L. (2020) 'The Evolution of Flight and Wing Shape in *Heliconius* Butterflies'. University of York.

Queste, L. M. (2015) *Characterising reproductive barriers between three closely related* Heliconius *butterfly taxa.* University of York.

Rabosky, D. L. (2017) 'Phylogenetic tests for evolutionary innovation: the problematic link between key innovations and exceptional diversification', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society Publishing, 372(1735), p. 20160417. doi: 10.1098/rstb.2016.0417.

Rafaeli, A. and Soroker, V. (1989) 'Influence of diel rhythm and brain hormone on pheromone production in two lepidopteran species', *Journal of Chemical Ecology*. Kluwer Academic Publishers-Plenum Publishers, 15(2), pp. 447–455. doi: 10.1007/BF01014691.

Rahmani, R. *et al.* (2020) 'Identification and Synthesis of Putative Pheromone Components of the Threatened Salt Marsh Bagworm Moth, *Whittleia retiella* (Lepidoptera: Psychidae)', *Journal of Chemical Ecology*. Springer, 46(2), pp. 115–127. doi: 10.1007/s10886-020-01145-x.

Raina, A. K., Klun, J. A. and Stadelbacher, E. A. (1986) 'Diel Periodicity and Effect of Age and Mating on Female Sex Pheromone Titer in *Heliothis zea* (Lepidoptera: Noctuidae)1', *Annals of the Entomological Society of America*. Oxford University Press (OUP), 79(1), pp. 128–131. doi: 10.1093/aesa/79.1.128.

Rastas, P. (2017) 'Lep-MAP3: Robust linkage mapping even for low-coverage whole genome sequencing data', *Bioinformatics*. Oxford University Press, 33(23), pp. 3726–3732. doi: 10.1093/bioinformatics/btx494.

Ratcliffe, L. M. and Grant, P. R. (1985) 'Species recognition in Darwin's finches (*Geospiza*, Gould). III.

Male responses to playback of different song types, dialects and heterospecific songs', *Animal Behaviour*. Academic Press, 33(1), pp. 290–307. doi: 10.1016/S0003-3472(85)80143-3.

Reddy, G. V. P. and Guerrero, A. (2004) 'Interactions of insect pheromones and plant semiochemicals', *Trends in Plant Science*, pp. 253–261. doi: 10.1016/j.tplants.2004.03.009.

Reed, J. R. *et al.* (1994) 'Unusual mechanism of hydrocarbon formation in the housefly: Cytochrome P450 converts aldehyde to the sex pheromone component (Z)-9-tricosene and CO2', *Proceedings of the National Academy of Sciences of the United States of America*, 91(21), pp. 10000–10004. doi: 10.1073/pnas.91.21.10000.

Reed, J. R. *et al.* (1995) 'Proposed Mechanism for the Cytochrome P450-Catalyzed Conversion of Aldehydes to Hydrocarbons in the House Fly, *Musca domestica*', *Biochemistry*, 34(49), pp. 16221–16227. doi: 10.1021/bi00049a038.

Reed, R. D. *et al.* (2011) 'optix Drives the Repeated Convergent Evolution of Butterfly Wing Pattern Mimicry', *Science*, 333(6046), pp. 1137–1141.

Regnier, F. E. and Law, J. H. (1968) 'Insect pheromones.', *Journal of Lipid Research*. American Society for Biochemistry and Molecular Biology, pp. 541–551. doi: 10.2307/2403157.

Revell, L. (2017) *Phylogenetic Tools for Comparative Biology: Pearson correlation with phylogenetic data*. Available at: http://blog.phytools.org/2017/08/pearson-correlation-with-phylogenetic.html (Accessed: 22 February 2021).

Revell, L. J. (2012) 'phytools: an R package for phylogenetic comparative biology (and other things)', *Methods in Ecology and Evolution*. John Wiley & Sons, Ltd (10.1111), 3(2), pp. 217–223. doi: 10.1111/j.2041-210X.2011.00169.x.

Revell, L. J., Harmon, L. J. and Collar, D. C. (2008) 'Phylogenetic Signal, Evolutionary Process, and Rate', *Systematic Biology*. Edited by T. Oakley. Narnia, 57(4), pp. 591–601. doi: 10.1080/10635150802302427.

Rheindt, F. E., Grafe, T. U. and Abouheif, E. (2004) 'Rapidly evolving traits and the comparative method: how important is testing for phylogenetic signal?', *Evolutionary Ecology Research*, 6, pp. 377–396.

Rieseberg, L. H. and Blackman, B. K. (2010) 'Speciation genes in plants', *Annals of Botany*, 106(3), pp. 439–455. doi: 10.1093/aob/mcq126.

Rochat, D. *et al.* (2000) 'Role of solid-phase microextraction in the identification of highly volatile pheromones of two Rhinoceros beetles *Scapanes australis* and *Strategus aloeus* (Coleoptera, Scarabaeidae, Dynastinae)', *Journal of Chromatography A*, 885(1–2), pp. 433–444. doi: 10.1016/S0021-9673(99)01049-3.

Rodríguez, S. *et al.* (2004) 'Expression and evolution of Δ9 and Δ11 desaturase genes in the moth *Spodoptera littoralis*', *Insect Biochemistry and Molecular Biology*, 34(12), pp. 1315–1328. doi: 10.1016/j.ibmb.2004.09.003.

Roelofs, W. L. and Rooney, A. P. (2003) 'Molecular genetics and evolution of pheromone biosynthesis in Lepidoptera', *Proceedings of the National Academy of Sciences*, 100(Supplement 2), pp. 14599–14599. doi: 10.1073/pnas.1233767100.

Rollmann, S. M., Houck, L. D. and Feldhoff, R. C. (2000) 'Population variation in salamander courtship pheromones', *Journal of Chemical Ecology*, 26(12), pp. 2713–2724. doi: 10.1023/A:1026429508058.

Ross, G. N. *et al.* (2001) 'Novel chemistry of abdominal defensive glands of nymphalid butterfly *Agraulis vanillae*', *Journal of Chemical Ecology*. Springer, 27(6), pp. 1219–1228. doi:

10.1023/A:1010372114144.

Ross, P. and Crews, D. (1977) 'Influence of the seminal plug on mating behaviour in the garter snake', *Nature*, pp. 344–345. doi: 10.1038/267344a0.

Ross, P. and Crews, D. (1978) 'Stimuli influencing mating behavior in the garter snake, *Thamnophis radix*', *Behavioral Ecology and Sociobiology*. Springer-Verlag, 4(2), pp. 133–142. doi: 10.1007/BF00354976.

Rosser, N. *et al.* (2012) 'Testing historical explanations for gradients in species richness in heliconiine butterflies of tropical America', *Biological Journal of the Linnean Society*, 105, pp. 479–497. doi: 10.1111/j.1095-8312.2011.01814.x.

Rosser, N. *et al.* (2015) 'Extensive range overlap between Heliconiine sister species: Evidence for sympatric speciation in butterflies?', *BMC Evolutionary Biology*, 15(1). doi: 10.1186/s12862-015-0420-3.

Rosser, N. *et al.* (2019) 'Geographic contrasts between pre- and postzygotic barriers are consistent with reinforcement in *Heliconius* butterflies', *Evolution*, 73(9), pp. 1821–1838. doi: 10.1111/evo.13804.

Rudie, K. A. (2015) *Effect of Captivity on Chemical Profiles of the Sexual Attractiveness Pheromone of the Red-Sided Garter Snake, <i>Thamnophis sirtalis parietalis*.

Rundle, H. D. and Nosil, P. (2005) 'Ecological speciation', *Ecology Letters*, 8, pp. 336–352. doi: 10.1111/j.1461-0248.2004.00715.x.

Ruther, J. *et al.* (2014) 'Pheromone Diversification and Age-Dependent Behavioural Plasticity Decrease Interspecific Mating Costs in *Nasonia*', *PLoS ONE*. Edited by W. J. Etges. Public Library of Science, 9(2), p. e89214. doi: 10.1371/journal.pone.0089214.

Salazar, C. *et al.* (2010) 'Genetic evidence for hybrid trait speciation in *Heliconius* butterflies', *PLoS Genetics*. Public Library of Science, 6(4). doi: 10.1371/journal.pgen.1000930.

Schaal, B. *et al.* (2003) 'Chemical and behavioural characterization of the rabbit mammary pheromone', *Nature*. Nature Publishing Group, 424(6944), pp. 68–72. doi: 10.1038/nature01739.

Schal, C., Sevala, V. and Cardé, R. T. (1998) 'Novel and highly specific transport of a volatile sex pheromone by hemolymph lipophorin in moths', *Naturwissenschaften*. Springer, 85(7), pp. 339–342. doi: 10.1007/s001140050511.

Schluter, D. (2001) 'Ecology and the origin of species', *Trends in Ecology and Evolution*. Elsevier Current Trends, pp. 372–380. doi: 10.1016/S0169-5347(01)02198-X.

Schulz, S. *et al.* (2007) 'Macrolides from the scent glands of the tropical butterflies *Heliconius cydno* and *Heliconius pachinus*', *Organic and Biomolecular Chemistry*. The Royal Society of Chemistry, 5(21), pp. 3434–3441. doi: 10.1039/b710284d.

Schulz, S. *et al.* (2008) 'An antiaphrodisiac in *Heliconius melpomene* butterflies', *Journal of Chemical Ecology*, 34(1), pp. 82–93. doi: 10.1007/s10886-007-9393-z.

Schulz, S., Boppre, M. and Vane-Wright, R. I. (1993) 'Specific mixtures of secretions from male scent organs of African milkweed butterflies (Danainae)', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 342(1300), pp. 161–181. doi: 10.1098/rstb.1993.0144.

Schulz, S., Yildizhan, S. and van Loon, J. J. A. (2011) 'The Biosynthesis of Hexahydrofarnesylacetone in the Butterfly *Pieris brassicae*', *Journal of Chemical Ecology*. Springer-Verlag, 37(4), pp. 360–363. doi: 10.1007/s10886-011-9939-y.

Schwinn, K. *et al.* (2006) 'A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the Genus *Antirrhinum*', *Plant Cell*. American Society of Plant Biologists, 18(4), pp. 831–851. doi: 10.1105/tpc.105.039255.

Scott, D. (1986) 'Sexual mimicry regulates the attractiveness of mated *Drosophila melanogaster* females', *Proceedings of the National Academy of Sciences of the United States of America*, 83(21), pp. 8429–8433. doi: 10.1073/pnas.83.21.8429.

Sculfort, O. *et al.* (2021) 'Assessing the Role of Developmental and Environmental Factors in Chemical Defence Variation in Heliconiini Butterflies', *Journal of Chemical Ecology*. Springer, 47(6), pp. 577–587. doi: 10.1007/s10886-021-01278-7.

Seddon, N. (2005) 'Ecological adaptation and species recognition drives vocal evolution in neotropical suboscine birds', *Evolution*. Society for the Study of Evolution, 59(1), pp. 200–215. doi: 10.1111/j.0014-3820.2005.tb00906.x.

Seehausen, O. *et al.* (2008) 'Speciation through sensory drive in cichlid fish', *Nature*. Nature Publishing Group, 455(7213), pp. 620–626. doi: 10.1038/nature07285.

Seehausen, O. and Van Alphen, J. J. M. (1998) 'The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex)', *Behavioral Ecology and Sociobiology*, 42(1), pp. 1–8. doi: 10.1007/s002650050405.

Sekimoto, H. (2005) 'Plant Sex Pheromones', *Vitamins and Hormones*. Vitam Horm, pp. 457–478. doi: 10.1016/S0083-6729(05)72013-6.

Servedio, M. R. *et al.* (2011) 'Magic traits in speciation: "magic" but not rare?', *Trends in Ecology & Evolution*, 26(8), pp. 389–397. doi: 10.1016/j.tree.2011.04.005.

Servedio, M. R. and Noor, M. A. F. (2003) 'The Role of Reinforcement in Speciation: Theory and Data', *Annual Review of Ecology, Evolution, and Systematics*, 34(1), pp. 339–364. doi: 10.1146/annurev.ecolsys.34.011802.132412.

Shahandeh, M. P., Pischedda, A. and Turner, T. L. (2018) 'Male mate choice via cuticular hydrocarbon pheromones drives reproductive isolation between *Drosophila* species', *Evolution*. Society for the Study of Evolution, 72(1), pp. 123–135. doi: 10.1111/evo.13389.

Shaw, K. L. and Lesnick, S. C. (2009) 'Genomic linkage of male song and female acoustic preference QTL underlying a rapid species radiation', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 106(24), pp. 9737–9742. doi: 10.1073/pnas.0900229106.

Shaw, K. L. and Mullen, S. P. (2014) 'Speciation Continuum', *Journal of Heredity*. Oxford University Press, 105(S1), pp. 741–742. doi: 10.1093/jhered/esu060.

Shockey, J. M., Fulda, M. S. and Browse, J. A. (2002) '*Arabidopsis* contains nine long-chain acyl-coenzyme A synthetase genes that participate in fatty acid and glycerolipid metabolism', *Plant Physiology*. American Society of Plant Biologists, 129(4), pp. 1710–1722. doi: 10.1104/pp.003269.

Shu, S. and Jones, R. L. (1993) 'Evidence for a multicomponent sex pheromone in *Eriborus terebrans* (Gravenhorst) (HYM.: Ichneumonidae), a larval parasitoid of the European corn borer', *Journal of Chemical Ecology*. Kluwer Academic Publishers-Plenum Publishers, 19(11), pp. 2563–2576. doi: 10.1007/BF00980691.

Silva-Brandão, K. L. *et al.* (2008) 'Phylogenetic relationships of butterflies of the tribe Acraeini (Lepidoptera, Nymphalidae, Heliconiinae) and the evolution of host plant use', *Molecular Phylogenetics and Evolution*, 46(2), pp. 515–531. doi: 10.1016/j.ympev.2007.11.024.

Silva, W. D. *et al.* (2017) 'Aggregation-Sex pheromones and likely pheromones of 11 South American cerambycid beetles, and Partitioning of pheromone channels', *Frontiers in Ecology and Evolution*. Frontiers Media S. A, 5(AUG). doi: 10.3389/fevo.2017.00101.

Silvestro, D., Zizka, G. and Schulte, K. (2014) 'Disentangling the effects of key innovations on the diversification of Bromelioideae (Bromeliaceae)', *Evolution*, 68(1), pp. 163–175. doi: 10.1111/evo.12236.

Simmons, L. W. (2001) *Sperm competition and its evolutionary consequences in the insects*. Princeton University Press.

Siva-Jothy, M. T. *et al.* (1996) 'Short- and long-term sperm precedence in the beetle *Tenebrio molitor*: A test of the "adaptive sperm removal" hypothesis', *Physiological Entomology*. Blackwell Publishing Ltd, 21(4), pp. 313–316. doi: 10.1111/j.1365-3032.1996.tb00871.x.

Slotman, M. A. *et al.* (2006) 'Reduced recombination rate and genetic differentiation between the M and S forms of *Anopheles gambiae* s.s.', *Genetics*. Oxford University Press, 174(4), pp. 2081–2093. doi: 10.1534/genetics.106.059949.

Smadja, C. and Butlin, R. K. (2009) 'On the scent of speciation: the chemosensory system and its role in premating isolation', *Heredity*. Springer International Publishing, 102(1), pp. 77–97. doi: 10.1038/hdy.2008.55.

Smadja, C. M. and Butlin, R. K. (2011) 'A framework for comparing processes of speciation in the presence of gene flow', *Molecular Ecology*. Mol Ecol, pp. 5123–5140. doi: 10.1111/j.1365-294X.2011.05350.x.

Smith, D. T. *et al.* (2017) 'Sexual conflict over remating interval is modulated by the sex peptide pathway', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 284(1850). doi: 10.1098/rspb.2016.2394.

Snell, T. W. (1998) 'Chemical ecology of rotifers', in *Hydrobiologia*. Springer Netherlands, pp. 267–276. doi: 10.1023/a:1017087003334.

Sobel, J. M. *et al.* (2010) 'The biology of speciation', *Evolution*, pp. 295–315. doi: 10.1111/j.1558-5646.2009.00877.x.

Solensky, M. J. and Oberhauser, K. S. (2009) 'Sperm precedence in monarch butterflies (*Danaus plexippus*)', *Behavioral Ecology*, 20(2), pp. 328–334. doi: 10.1093/beheco/arp003.

South, S. H. *et al.* (2011) 'Male cockroaches prefer a high carbohydrate diet that makes them more attractive to females: Implications for the study of condition dependence', *Evolution*. John Wiley & Sons, Ltd, 65(6), pp. 1594–1606. doi: 10.1111/j.1558-5646.2011.01233.x.

Southcott, L. and Kronforst, M. R. (2018) 'Female mate choice is a reproductive isolating barrier in *Heliconius* butterflies', *Ethology*. Blackwell Publishing Ltd, 124(12), pp. 862–869. doi: 10.1111/eth.12818.

Sparkman, O. D., Penton, Z. and Kitson, F. G. (2011) *Gas chromatography and mass spectrometry: a practical guide*. Academic Press.

Squires, Z. E. *et al.* (2015) 'Last male sperm precedence in a polygamous squid', *Biological Journal of the Linnean Society*. Blackwell Publishing Ltd, 116(2), pp. 277–287. doi: 10.1111/bij.12590.

Stacey, N. (2003) 'Hormones, pheromones and reproductive behavior', *Fish Physiology and Biochemistry*, 28(1–4), pp. 229–235. doi: 10.1023/B:FISH.0000030540.99732.2c.

Stanley, J. *et al.* (2018) 'Evidence of male pheromone in *Conogethes punctiferalis* (Lepidoptera:

Pyralidae)', *Journal of Entomological Science*, 53(4), pp. 455–466. doi: 10.18474/JES17-98.1.

Stein, S. E. (1999) 'An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data', *Journal of the American Society for Mass Spectrometry*. No longer published by Elsevier, 10(8), pp. 770–781. doi: 10.1016/S1044-0305(99)00047-1.

Steiner, S., Steidle, J. L. M. and Ruther, J. (2005) 'Female sex pheromone in immature insect males - A case of pre-emergence chemical mimicry?', *Behavioral Ecology and Sociobiology*. Springer, 58(2), pp. 111–120. doi: 10.1007/s00265-005-0930-x.

Sun, X. L. *et al.* (2010) 'The tea weevil, *Myllocerinus aurolineatus*, is attracted to volatiles induced by conspecifics', *Journal of Chemical Ecology*. Springer, 36(4), pp. 388–395. doi: 10.1007/s10886-010-9771-9.

Supple, M. A. *et al.* (2015) 'Divergence with gene flow across a speciation continuum of *Heliconius* butterflies', *BMC Evolutionary Biology*. BioMed Central, 15(1), p. 204. doi: 10.1186/s12862-015-0486-y.

Symonds, Matthew R. E. and Elgar, M. A. (2004) 'Species overlap, speciation and the evolution of aggregation pheromones in bark beetles', *Ecology Letters*. John Wiley & Sons, Ltd (10.1111), 7(3), pp. 202–212. doi: 10.1111/j.1461-0248.2004.00571.x.

Symonds, Matthew R.E. and Elgar, M. A. (2004) 'The mode of pheromone evolution: Evidence from bark beetles', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 271(1541), pp. 839–846. doi: 10.1098/rspb.2003.2647.

Symonds, M. R. E. and Elgar, M. A. (2008) 'The evolution of pheromone diversity', *Trends in Ecology and Evolution*. Elsevier Current Trends, 23(4), pp. 220–228. doi: 10.1016/j.tree.2007.11.009.

Symonds, M.R.E. and Gitau-Clarke, C. W. (2016) 'The Evolution of Aggregation Pheromone Diversity in Bark Beetles', *Advances in Insect Physiology*. Academic Press, 50, pp. 195–234. doi: 10.1016/BS.AIIP.2015.12.003.

Symonds, M. R.E. and Gitau-Clarke, C. W. (2016) 'The Evolution of Aggregation Pheromone Diversity in Bark Beetles', in *Advances in Insect Physiology*. Academic Press Inc., pp. 195–234. doi: 10.1016/bs.aiip.2015.12.003.

Taga, M. E. and Bassler, B. L. (2003) 'Chemical communication among bacteria', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 100(24), pp. 14549–14554. doi: 10.1073/pnas.1934514100.

Tang, W. *et al.* (2018) 'Estimation of QTL heritability based on pooled sequencing data', *Bioinformatics*. Edited by O. Stegle. Oxford University Press, 34(6), pp. 978–984. doi: 10.1093/bioinformatics/btx703.

Thomas, G. H. and Freckleton, R. P. (2012) 'MOTMOT: models of trait macroevolution on trees', *Methods in Ecology and Evolution*. John Wiley & Sons, Ltd, 3(1), pp. 145–151. doi: 10.1111/j.2041-210X.2011.00132.x.

Thurman, T. J. *et al.* (2018) 'Facultative pupal mating in *Heliconius erato* : Implications for mate choice, female preference, and speciation', *Ecology and Evolution*. John Wiley and Sons Ltd, 8(3), pp. 1882–1889. doi: 10.1002/ece3.3624.

Tillman, J. A. *et al.* (1999) 'Insect pheromones - An overview of biosynthesis and endocrine regulation', *Insect Biochemistry and Molecular Biology*, pp. 481–514. doi: 10.1016/S0965-1748(99)00016-8.

Ting, C. T. *et al.* (1998) 'A rapidly evolving homeobox at the site of a hybrid sterility gene', *Science*. American Association for the Advancement of Science, 282(5393), pp. 1501–1504. doi: 10.1126/science.282.5393.1501.

Ting, J. H., Kelly, L. S. and Snell, T. W. (2000) 'Identification of sex, age and species-specific proteins on the surface of the harpacticoid copepod *Tigriopus japonicus*', *Marine Biology*, 137(1), pp. 31–37. doi: 10.1007/s002270000320.

Tomalski, M. D. *et al.* (1987) 'Chemistry and functions of exocrine secretions of the ants *Tapinoma melanocephalum* and *T. erraticum*', *Journal of Chemical Ecology*. Kluwer Academic Publishers-Plenum Publishers, 13(2), pp. 253–263. doi: 10.1007/BF01025886.

Tompkins, L. and Hall, J. C. (1981) '*Drosophila* males produce a pheromone which inhibits courtship', *Zeitschrift fur Naturforschung - Section C Journal of Biosciences*, 36(7–8), pp. 694–696. doi: 10.1515/znc-1981-7-832.

Toure, M. W. *et al.* (2020) 'Heliconiini butterflies can learn time-dependent reward associations: Evidence of time learning in butterflies', *Biology Letters*. Royal Society Publishing, 16(9). doi: 10.1098/rsbl.2020.0424.

Turner, T. L., Hahn, M. W. and Nuzhdin, S. V (2005) 'Genomic Islands of Speciation in *Anopheles gambiae*', *PLoS Biology*. Edited by N. Barton. Public Library of Science, 3(9), p. e285. doi: 10.1371/journal.pbio.0030285.

Vanderwel, D. (1994) 'Factors affecting pheromone production in beetles', *Archives of Insect Biochemistry and Physiology*. John Wiley & Sons, Ltd, 25(4), pp. 347–362. doi: 10.1002/arch.940250409.

Vane-Wright, R. I. and Boppre, M. (1993) 'Visual and chemical signalling in butterflies: functional and phylogenetic perspectives', *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. The Royal SocietyLondon, 340(1292), pp. 197–205. doi: 10.1098/rstb.1993.0058.

Vanjari, S. *et al.* (2015) 'Male sex pheromone components in the butterfly *Heliconius melpomene*.' doi: 10.1101/033506.

Vas, G. and Vékey, K. (2004) 'Solid-phase microextraction: a powerful sample preparation tool prior to mass spectrometric analysis', *Journal of Mass Spectrometry*. John Wiley & Sons, Ltd, 39(3), pp. 233–254. doi: 10.1002/jms.606.

Via, S. (2001) 'Sympatric speciation in animals: The ugly duckling grows up', *Trends in Ecology and Evolution*. Trends Ecol Evol, pp. 381–390. doi: 10.1016/S0169-5347(01)02188-7.

Via, Sara *et al.* (2012) 'Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow.', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. The Royal Society, 367(1587), pp. 451–60. doi: 10.1098/rstb.2011.0260.

Vincenten, N. *et al.* (2015) 'The kinetochore prevents centromere-proximal crossover recombination during meiosis', *eLife*. eLife Sciences Publications Ltd, 4(DECEMBER2015). doi: 10.7554/eLife.10850.

Wagenmakers, E.-J. and Farrell, S. (2004) 'AIC model selection using Akaike weights', *Psychonomic Bulletin & Review*. Springer-Verlag, 11(1), pp. 192–196. doi: 10.3758/BF03206482.

Walsh, H. E. *et al.* (1999) 'Polytomies and the Power of Phylogenetic Inference', *Evolution*. JSTOR, 53(3), p. 932. doi: 10.2307/2640732.

Walters, J. R. *et al.* (2012) 'Evaluating female remating rates in light of spermatophore degradation in *Heliconius* butterflies: Pupal-mating monandry versus adult-mating polyandry', *Ecological Entomology*, 37(4), pp. 257–268. doi: 10.1111/j.1365-2311.2012.01360.x.

Wang, H.-L. *et al.* (2014) 'Identification and Biosynthesis of Novel Male Specific Esters in the Wings of the Tropical Butterfly, *Bicyclus martius sanaos*', *Journal of Chemical Ecology*, 40, pp. 549–559. doi: 10.1007/s10886-014-0452-y.

Watson, G. J., Hamilton, K. M. and Tuffnail, W. E. (2005) 'Chemical alarm signalling in the polychaete *Nereis (Neanthes) virens* (Sars) (Annelida: Polychaeta)', *Animal Behaviour*, 70(5), pp. 1125–1132. doi: 10.1016/j.anbehav.2005.03.011.

Watson, P. J., Arnqvist, G. and Stallmann, R. R. (1998) 'Sexual conflict and the energetic costs of mating and mate choice in water striders', *American Naturalist*, 151(1), pp. 46–58. doi: 10.1086/286101.

Weber, M. G. *et al.* (2016) 'Macroevolution of perfume signalling in orchid bees', *Ecology Letters*, 19(11), pp. 1314–1323. doi: 10.1111/ele.12667.

Wei, T. and Simko, V. (2017) 'Package "corrplot" for R: Visualization of a Correlation Matrix', *CRAN*.

Westerman, E. L. *et al.* (2018) 'Aristaless Controls Butterfly Wing Color Variation Used in Mimicry and Mate Choice', *Current Biology*. Cell Press, 28(21), pp. 3469-3474.e4. doi: 10.1016/j.cub.2018.08.051.

Wheeler, Q. and Rudolf, M. (2000) *Species Concepts and Phylogenetic Theory: A Debate*. New York: Columbia University Press.

Wicker-Thomas, C. (2007) 'Pheromonal communication involved in courtship behavior in Diptera', *Journal of Insect Physiology*. Pergamon, pp. 1089–1100. doi: 10.1016/j.jinsphys.2007.07.003.

Wicker-Thomas, C. (2011) 'Evolution of insect pheromones and their role in reproductive isolation and speciation', *Annales de la Societe Entomologique de France*, 47(1–2), pp. 55–62. doi: 10.1080/00379271.2011.10697696.

Wigby, S. and Chapman, T. (2005) 'Sex peptide causes mating costs in female *Drosophila melanogaster*', *Current Biology*. Cell Press, 15(4), pp. 316–321. doi: 10.1016/j.cub.2005.01.051.

Wiley, C. and Shaw, K. L. (2010) 'Multiple genetic linkages between female preference and male signal in rapidly speciating hawaiian crickets', *Evolution*. John Wiley & Sons, Ltd, 64(8), pp. 2238–2245. doi: 10.1111/j.1558-5646.2010.01007.x.

Wiley, E. O. and Mayden, R. L. (2000) 'The Evolutionary Species Concept', in Wheeler, Q. and Rudolf, M. (eds) *Species Concepts and Phylogenetic Theory: A Debate*, pp. 70–92.

Williams, P. *et al.* (2007) 'Look who's talking: Communication and quorum sensing in the bacterial world', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, pp. 1119–1134. doi: 10.1098/rstb.2007.2039.

Wilmes, A. J. *et al.* (2012) 'Conspecific pheromone trailing and pheromone trail longevity in the African Colubrid Boaedon fuliginosus', *African Journal of Herpetology*, 61(2), pp. 159–168. doi: 10.1080/21564574.2012.713397.

Wirth, R., Muscholl, A. and Wanner, G. (1996) 'The role of pheromones in bacterial interactions', *Trends in Microbiology*. Elsevier Ltd, pp. 96–103. doi: 10.1016/0966-842X(96)81525-3.

Witzgall, P., Kirsch, P. and Cork, A. (2010) 'Sex pheromones and their impact on pest management', *Journal of Chemical Ecology*, 36(1), pp. 80–100. doi: 10.1007/s10886-009-9737-y.

Wu, C. I. (2001) 'The genic view of the process of speciation', *Journal of Evolutionary Biology*. John Wiley & Sons, Ltd, pp. 851–865. doi: 10.1046/j.1420-9101.2001.00335.x.

Wu, C. I. and Ting, C. T. (2004) 'Genes and speciation', *Nature Reviews Genetics*, pp. 114–122. doi: 10.1038/nrg1269.

Würf, J. *et al.* (2020) 'Cuticular Hydrocarbons as Contact Sex Pheromone in the Parasitoid Wasp *Urolepis rufipes*', *Frontiers in Ecology and Evolution*. Frontiers Media S.A., 8, p. 180. doi: 10.3389/fevo.2020.00180.

Wyatt, T. (2009) 'Pheromones and other Chemical Communication in Animals', in *Encyclopedia of Neuroscience*. Elsevier Ltd, pp. 611–616. doi: 10.1016/B978-008045046-9.01825-8.

Xu, J. and Wang, Q. (2010) 'Mechanisms of last male precedence in a moth: Sperm displacement at ejaculation and storage sites', *Behavioral Ecology*, 21(4), pp. 714–721. doi: 10.1093/beheco/arq044.

Yan, S. *et al.* (2013) 'Sperm storage and sperm competition in the *Helicoverpa armigera* (Lepidoptera: Noctuidae)', *Journal of Economic Entomology*. Oxford University Press (OUP), 106(2), pp. 708–715. doi: 10.1603/EC12402.

Yew, J. Y. and Chung, H. (2015) 'Insect pheromones: An overview of function, form, and discovery', *Progress in Lipid Research*, 59, pp. 88–105. doi: 10.1016/j.plipres.2015.06.001.

Yoder, J. B. *et al.* (2010) 'Ecological opportunity and the origin of adaptive radiations', *Journal of Evolutionary Biology*. John Wiley & Sons, Ltd, pp. 1581–1596. doi: 10.1111/j.1420-9101.2010.02029.x.

Young, F. J. and Montgomery, S. H. (2020) 'Pollen feeding in *Heliconius* butterflies: the singular evolution of an adaptive suite', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society Publishing. doi: 10.1098/rspb.2020.1304.

Zhang, D.-D. and Löfstedt, C. (2015) 'Moth pheromone receptors: gene sequences, function, and evolution', *Frontiers in Ecology and Evolution*. Frontiers Media S. A, 3(SEP), p. 105. doi: 10.3389/fevo.2015.00105.

Zhang, W. *et al.* (2016) 'Genome-wide introgression among distantly related *Heliconius* butterfly species.', *Genome biology*. BioMed Central, 17, p. 25. doi: 10.1186/s13059-016-0889-0.

Zhang, W. F. *et al.* (2022) 'Production and evolution pattern of "fruity smell" aggregation pheromones in genus *Drosophila*', *Journal of Systematics and Evolution*. John Wiley and Sons Inc, 60(1), pp. 208–219. doi: 10.1111/jse.12648.

Zimmermann, Y., Ramirez, S. R. and Eltz, T. (2009) 'Chemical niche differentiation among sympatric species of orchid bees', *Ecology*, 90(11), pp. 2994–3008. doi: 10.1890/08-1858.1.