9-22-2017 9:30 AM

The genetic and environmental basis for CHC biosynthesis in Drosophila

Heather KE Ward, *The University of Western Ontario*

Supervisor: Dr. Amanda Moehring, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biology

Abstract

Cuticular hydrocarbons (CHCs) are produced by insects and primarily used to prevent desiccation. In *Drosophila*, certain compounds have secondary roles as infochemicals that may act during courtship to influence mate choice. Certain CHCs may stimulate courtship with heterospecifics or act to repel conspecifics. The CHC profile produced by an individual is the result of the interaction between its genetic background and the environment, though the genes that underlie species differences in CHC production and how the environment can modulate the abundance of individual compounds within a species is not well known. Here, candidate gene *CG5946* was found to be involved in species differences in the production of 7,11-heptacosadiene and 7-tricosene in hybrids between *D. melanogaster* and *D. simulans*. In addition, diet, but not microbial content, was found to influence the proportion of long-chain CHCs produced by *D. melanogaster*. This study provides insight into the factors influencing CHC production in *Drosophila.*

Keywords

Drosophila, cuticular hydrocarbons, speciation, evolution, mate choice, behavioural isolation

Acknowledgments

I would first like to thank Dr. Amanda Moehring for all of her support and guidance throughout my time at Western. Her presence and input has grounded and guided me, and her deep knowledge of and passion for the study of *Drosophila*, as well as her enthusiasm for teaching, has inspired me to see beyond the fly. I will always be grateful for the lessons I have learned under her supervision, and I wish her all the best.

I am also incredibly grateful to Dr. Mark Bernards, who not only provided insight as an advisor but also helped me to run countless samples and remained calm when one machine after another broke down. I can't thank him enough for his help and patience during my time here.

Many thanks also to Dr. Jeremy McNeil, advisory committee member and entomologist-extraordinaire, whose encyclopaedic knowledge of insects and advice throughout this project has helped to keep looking in new directions.

Thank you also to all of my lab mates, past and present: Rachelle Kanippayoor, Ryan Calhoun, Tabashir Chowdhury, Caryn Dooner, Josh Isaacson, and especially Pria Mahabir and Jalina Bielaska Da Silva, my fellow master's students. Without the support of my labmates, all of my flies would have died during week one and I'd be left with a purely theoretical work. Having Pria and Jalina as a support group and completing our projects together has been all that I could have asked for from a lab. I wish them, and the rest of the Moehring lab, all future success as they pursue science in whatever form it takes them; I know they will do great things.

Finally, thanks to my family for listening to me ramble about hydrocarbons and metabolism and other such things that are very likely wholly uninteresting to anyone who is not really, really into flies. Thanks also to my partners Chris and Spencer, who always make me want to reach higher and do more.

Table of Contents

List of Tables

List of Figures

Chapter 1

1 General Introduction

1.1 Species

The question of what constitutes a species has been a central theme in biology since its inception. Historically, taxonomists have separated individuals into groups based on traits such as morphological similarity, but the advent of modern methods such as DNA sequencing has seen the decline of such limited classification schemes. The use of morphology to separate individuals into species can result in errors on both ends of the scale: evolutionarily distinct but morphologically similar organisms can be classed together due to convergent evolution, and members of the same species may be separated into different taxa if males and females are sexually dimorphic or if individuals are collected at different life stages (phase polymorphism - as in the case of locusts and grasshoppers) (Pener and Yerushalmi, 1998). Nowadays, tens of different species concepts exist, each of which can yield wildly varying counts of apparent species when looking at the same organisms. The most basic definition of a species is that it is the largest non-arbitrary evolving unit above the level of the individual, but what this definition involves in a biological setting is the topic of heated debate (reviewed in: Coyne and Orr, 2004).

Darwin himself expressed doubts about whether or not species are things that actually exist in nature or whether they are simply concepts that act as useful approximations, stating:

> *"...I look at the term species, as one arbitrarily given for the sake of convenience to a set of individuals closely resembling each other..."*
> (Darwin, 1859)

It is clear that we see patterns in nature that we naturally tend to group together, however the basis for this classification is difficult to resolve. Darwin recognized that the forms observed in nature are continuous, not discrete, and so cannot be easily separated into distinct species groups using any one criterion (Darwin, 1859). DNA sequencing has

confirmed that the continuous nature of species is also true at the genetic level – groups can have large genotypic and phenotypic variation, and yet be widely accepted to be a single species. Indeed, the process of evolution requires that populations exhibit variation. If species develop gradually via evolution, how then can we choose where to delimit where one species ends and another begins if a continuum between different forms exists? As Leclerc states:

> *"In general, the more one increases the number of divisions, in the case of natural products, the nearer one comes to the truth; since in reality individuals alone exist in nature, while genera, orders, classes exist only in our imagination."* (Leclerc, G-L., 1766)

The more variables we include in our species definition, the closer we come to classifying individuals as species. While this definition may come closest to the truth, it is not as functionally useful as other species definitions. The species concepts in use today are largely delimited using the trait or traits preferred by the particular scientist performing the classification, but the traits useful for one clade may prove entirely irrelevant in another, as is the case for many species concepts between sexually and asexually reproducing organisms. It is possible that a functional, unifying concept of what constitutes a species cannot be found that will successfully describe all of the forms of life observed in nature.

Although it appears that any species groupings that we make are based on artificial divisions, the adoption of operational species concepts can be useful in asking questions about how lineages evolve and diverge over time. Accepting this, we can make use of species concepts that separate species based on a diverse set of traits to attempt to gain insight into which methods produce the most consistent results and what this means in terms of how speciation occurs. By examining the diversity of life that we observe today in the context of these different species concepts, we can hopefully begin to understand the factors that lead to divergence in nature, and how the different groups that we call species are able to form and remain distinct over time.

1.2 The biological species concept

One of the most widely used species concepts in sexually reproducing taxa is the biological species concept (BSC). The modern version of this concept was introduced by Ernst Mayr (1942), and states that *species are groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups*. The BSC uses the traits involved in reproduction and mate choice to delineate what makes a species distinct from other such groups. This species concept is useful in sexually reproducing taxa, as it can differentiate species living in sympatry regardless of whether they are reproductively isolated due to behavior, genetic incompatibility, or some other factor or a combination of factors. This means that even if two individuals could produce viable offspring if forced to mate in an artificial environment, if they fail to reproduce in nature for any reason and there is no or low observed gene flow between the groups, they can be classified as distinct species.

Although the BSC can be very useful, it also has limitations that prevent it from being used as a universal species concept. First among these is the fact that the BSC cannot be used to distinguish asexual or clonally reproducing taxa, or those groups that switch between sexual and asexual forms of reproduction. In addition, it is not possible to use the BSC to determine whether geographically isolated populations are reproductively isolated, or whether they would experience gene flow if brought into contact with one another (reviewed in: Coyne and Orr, 2004). Some versions of the definition include geographic isolation as a reproductive isolating barrier, where others treat this interpretation as too simplistic: shouldn't we call populations that have no genetic or behavioural differences a single species, regardless of the fact that they may inhabit a non-contiguous range (Sobel *et al.,* 2010)? Another problem with the BSC is its inability to distinguish fossil species, since it is not possible to determine whether extinct taxa could have experienced gene flow based on external morphology alone.

Despite the fact that the BSC has a number of practical drawbacks, its key feature, the fact that it distinguishes species based on reproductive capability, meshes well with the theory of evolution. Since evolution happens via changes in allele frequency within

populations over time (Curtis and Barnes, 1989) and is the force that produces the distinct groups that we recognize as species, it is clear that the ability of different populations to exchange genetic content via sexual reproduction is key in determining their evolutionary path. The use of a species concept that is based upon whether or not this gene flow can happen between populations is therefore particularly relevant in looking at species within an evolutionary context.

1.3 Reproductive isolating barriers

If species are to remain distinct when their range overlaps with that of another group (sympatric species), they must evolve a system of reproductive isolation that impedes gene flow between the groups. If such a system does not exist the populations will either merge into one species or one of the species (typically the rarer one) will go extinct (reviewed in: Rhymer and Simberloff, 1996). Reproductive barriers most often develop in allopatry and are usually the result of overall genetic displacement between groups, often as a side effect of adaptation to different local environments (reviewed in: Seehausen *et al.,* 2014). These barriers can be reinforced when species come into secondary contact, especially when hybrids suffer reduced fitness, since they will be less able than their non-hybrid peers to survive and reproduce (Oritz-Barrientos *et al.,* 2004). Alternatively, it has been suggested that reproductive barriers can evolve rapidly in sympatry, where a single species will split into two distinct groups due to forces such as assortative mating or divergent selection between the incipient species (Orr and Smith, 1998; Rieseberg *et al.,* 2002; reviewed in: Coyne and Orr, 2004). If directional selection acts in different directions on members of a single population, which can be due to differential degrees of preferences and traits due to underlying genetic variability, specialization on different resources, or some other factor, this will lead to divergent selection and can result in restricted gene flow between members of a population, eventually leading to speciation (Rieseberg *et al.,* 2002; reviewed in Servedio, 2016). The degree to which speciation occurs in sympatry is controversial, since the evolution of isolating barriers is slow and any gene flow rapidly purges the barriers that may begin to evolve between species that still live in close contact (reviewed in: Coyne and Orr, 2004; reviewed in: Servedio, 2016). Most models of sympatric speciation require linkage between the genes underlying preference for resources or traits and those involved

in reproductive isolation, such that divergent selection within a population is coupled with a reduced propensity or ability to mate so that gene flow is reduced as populations diverge (Friesen *et al*., 2007).

Reproductive isolating barriers are broadly classed into two groups: pre- and postzygotic barriers. Postzygotic barriers act following the formation of the gamete, and result in hybrids that are sterile or inviable (intrinsic postzygotic isolation), or those that cannot survive or find mates due to the interaction between the hybrid genetic background and the environment (extrinsic posyzygotic isolation) (reviewed in: Seehausen *et al.,* 2014). Prezygotic isolating barriers are those that act prior to the formation of a gamete and they typically act by preventing mating in the first place (reviewed in: Coyne and Orr, 2004), but there are also postmating, prezygotic isolating barriers, as is seen in copulatory and gametic isolation (Ludlow and Magurran, 2006). More common is behavioural isolation, in which some aspect of the courtship ritual of the individuals involved does not align with the expected signals such that the partner is not recognized as an appropriate mate (Dobzhansky, 1937; Martin and Hosken, 2003).

Despite the fact that individuals that are prezygotically isolated from one another may be genetically compatible, they do not experience gene flow in nature due to differences that prevent mating in the first place. The reinforcement of prezygotic isolating barriers occurs between species that are also postzygotically isolated, since investing resources into a hybrid that will not be able to survive or reproduce is costly and those individuals that are able to discriminate against heterospecifics will have a fitness advantage over those that are unable to discriminate (Dobzhansky, 1937).

1.4 Mate choice

The fitness of a sexually reproducing individual is dependent upon their ability to find an appropriate mate (Dobzhansky, 1937; reviewed in Andersson and Simmons, 2006). Mate choice is typically more important for the sex that invests more heavily in reproduction (Trivers, 1972; Burley, 1977). The sex that invests more heavily in reproduction is often the more ‘choosy’ partner, as choosing wrong is more costly to that partner, and this sex is

the one whose preferences are largely responsible for driving sexually selected traits (Trivers, 1972). In most cases this is the female of the species since females typically invest more resources into both gamete production (egg vs. sperm) and into raising offspring, though there are some examples of species in which the male is responsible for caring for and raising offspring and where males are the more discriminating sex, as in the buttonquail, *Turnix suscitator* (Starck, 1991).

Sexual selection often selects for traits that are actively detrimental to the survival ability of an animal (reviewed in: Kokko and Jennions, 2014). This is the case in songbirds such as the long-tailed widowbird, *Euplectes progne*. Females preferentially mate with males with longer tail feathers, however these feathers make it difficult for males to fly by adding significant drag, making them less capable hunters and easier prey (Craig, 1980).

Especially in those cases where postzygotic isolation exists between sympatric species, those individuals that are the most successful at preferentially mating with conspecifics will have the highest fitness (reviewed in: Coyne and Orr, 2004). Some of the most important factors in choosing a mate are ensuring that it is a) of the correct species, and b) a high quality mate. Choosing a mate of the wrong species is far more costly than choosing a low quality mate within the same species, as often this means that fecundity will be reduced to zero rather than simply being lowered (Kozak *et al.,* 2008). The signals that are involved in species recognition can overlap with, or be distinct from, those that are used to assess the quality of a conspecific individual. Individuals that are able to navigate both axes will produce the most offspring, which will inherit the genes underlying both species-specific preferences and traits.

Although courtship and mating rituals vary widely across taxa, recognition of potential mates ultimately stems from the ability of an animal to assess a conspecific's identity and quality by integrating multiple sensory cues. These can come in the form of auditory, visual, tactile, gustatory and/or olfactory signals (Horth, 2007; Griffith and Ejima, 2009; Ward and McLennan, 2009). Of the many signals used to recognize and assess potential mates, chemical signals in particular are widely used in a variety of animal

groups, since they can be sensed using multiple modalities (gustatory, olfactory, and/or tactile) (reviewed in: Johansson and Jones, 2007).

1.5 Cuticular hydrocarbons and chemical signaling in insects

Chemical communication is widely used by insects to transfer information both within and between species. The signals used in this communication range from small volatile molecules to long-chain hydrocarbon molecules that must be detected at close range (Jackson and Morgan, 1993). Chemical signals can induce behaviours such as aggregation, courtship, and aggression both within and between species (reviewed in: Ali and Morgan, 1990). In addition, chemical signals can provide information about nestmate status and caste, dominance, species, and sex (Howard and Blomquist, 2005; reviewed in: Blomquist and Bagnères, 2010).

Cuticular hydrocarbons (CHCs) are chemical compounds produced by all insects that are used for desiccation resistance and, in some species, chemical communication, including communication related to mating. CHCs are fatty-acid derivative molecules that form a waxy outer layer on the cuticle of the insect (Howard and Blomquist, 2005; reviewed in: Blomquist and Bagnères, 2010). They are produced in the oenocytes of the insect, which are specialized cells that are closely associated with the fat body and participate in fatty acid metabolism. These compounds are then transported to the cuticle (reviewed in: Blomquist and Bagnères, 2010), though the mechanisms involved in this transport are not well characterized. The major classes of CHCs are alkanes, alkenes, and methyl-branched hydrocarbons, and their primary role is to act as anti-desiccant molecules that prevent the insect from drying out and help to maintain an appropriate water balance (reviewed in: Blomquist and Bagnères, 2010). Long-chain saturated CHCs provide greater desiccation resistance than shorter chain and unsaturated or branched CHCs (Chung and Carroll, 2015).

In some insects, certain CHCs have developed a secondary role as infochemicals. CHCs that provide information within species are called pheromones while those that convey information between species are called allomones or kairomones, depending on

whether the emitter or the receiver benefits from the release of the compound (reviewed in: Ginzel, 2010). Insects display a CHC profile that is characteristic of their species that can be comprised of some 10-100+ unique compounds in various amounts (reviewed in: Blomquist and Bagnères, 2010). Some insects such as *Drosophila* show sexual dimorphism in the types or amounts of CHCs that they produce, and different profiles will be found on juvenile vs. adult insects (reviewed in: Bontonou and Wicker-Thomas, 2014).

In their role as infochemicals, CHCs largely provide information about the species, sex, nestmate status and caste of an individual (Howard and Blomquist, 2005; reviewed in: Blomquist and Bagnères, 2010). They are especially important in social insect societies, where the CHC profile of an individual can give nestmates information about the tasks a certain worker performs, as in the red harvester ant *Pogonomyrmex barbatus* (Greene and Gordon, 2003). In this species workers, foragers, patrollers, and nest-maintenance workers have distinct CHC profiles. When patrollers, whose job it is to search for food, return to the nest and interact with foragers, the foragers will exit the nest to retrieve food. Foragers are able to distinguish between returning patrollers and returning nest-maintenance workers based on their CHC profiles, and the return of nest-maintenance workers does not induce foraging activity (Greene and Gordon, 2003).

CHCs can also be used to act as dominance cues, as in the wasp species *Polistes dominulus* (Sledge *et al.*, 2001). These wasps have societies where a single alpha-female acts as a de-facto queen and is responsible for producing all offspring in the colony. When a group of females founds a new colony all females initially have identical CHC profiles; however, the alpha-female will become chemically distinct as soon as the first workers are born. If the alpha-female is removed from the colony, another female will establish herself as alpha, and her chemical profile will change accordingly. It has been suggested that using chemical signaling to convey dominance is an adaptation in these wasps, as behavioural dominance in the form of fighting for rank can come at high risk to the individuals involved (Sledge *et al.*, 2001).

Another widespread use for CHCs in insects is related to mimicry and deception. Many insects, rather than raising their own young, will lay eggs in a host's nest and use

chemical mimicry to convince the host to care for the alien young (Dettner and Liepert, 1994; reviewed in: Howard and Blomquist, 2005). Parasites can either mimic the CHC profile of the host by synthesizing the compounds involved *de novo* or they can acquire them via physical contact with their host insects (reviewed in: Howard and Blomquist, 2005). In another use of chemical mimicry, the parasitoid *Lysiphelebus cardui* attacks aphid colonies that are tended by the ant *Lasius niger*. *Lysiphelebus cardui* has evolved to mimic the CHC profile of the aphids such that the ants treat the parasitoid as though it were itself an aphid, allowing it to freely move amongst the ants and aphids (Liepert and Dettner, 1996).

One of the most important uses for CHCs is in their role as sex pheromones in some insects. Insects are able to detect the CHC profile of a potential mate and determine whether it is the same species, the proper sex, and whether or not it is sexually mature based on the relative abundance and types of CHCs that it produces (Howard and Blomquist, 2005; reviewed in: Blomquist and Bagnères, 2010). If the wrong profile is presented mating is unlikely to occur, making CHCs and the genes underlying their biosynthesis an important factor in both mate selection within a species, as well as prezygotic reproductive isolation and the maintenance of distinct species that live in sympatry.

1.6 Reproductive isolation in *Drosophila*

Drosophila melanogaster is part of the melanogaster subgroup of *Drosophila,* which includes eight other species (Ko *et al.*, 2003). These sibling species are the most genetically similar to *D. melanogaster*; however *D. melanogaster* exhibits intrinsic postzygotic isolation with all of its sibling species and forms inviable or sterile interspecies hybrids (Matute and Coyne, 2010; Cattani and Presgraves, 2012). Flies in the melanogaster subgroup also show high levels of prezygotic behavioural isolation (Ritchie *et al.*, 1999; McNabey, 2012; Matute, 2014). Courtship in *Drosophila* involves several steps at which species-specific cues are exchanged (detailed below in Section 1.7). Divergence in both the signals presented by flies and their perception by potential partners is responsible for the prezygotic isolation observed between *D. melanogaster* and its sibling species when they exist in sympatry (Savarit *et al.,* 1999; Billeter *et al.,* 2009).

The prezygotic barriers observed in *Drosophila* are strongly isolating, as evidenced by the fact that although the ranges of many of these species overlap, hybrids are found only very rarely in nature (Cattani and Presgraves, 2012). While the species will hybridize in the lab, they are reluctant to do so (reviewed in: Coyne and Orr, 2004). Since gene flow between these sibling species does not take place in nature, they represent complete/distinct species under the BSC.

1.7 The *Drosophila* courtship ritual

Courtship in *Drosophila* takes place in a series of well-defined steps. At each step, different sensory modalities are utilized to determine whether or not courtship and copulation will proceed (Griffith and Ejima, 2009). The female is ultimately responsible for deciding whether copulation will occur in most *Drosophila* species (Greenspan, 1995), although males courting heterospecific females may court at lower frequency than when courting conspecifics (Ellis and Carney, 2009).

During courtship, the male first approaches the female, orients himself towards her, then extends one of his forelegs to tap her on the abdomen (reviewed in: Spieth, 1974). The male has both gustatory and olfactory receptors on the last segment of his tarsus, allowing for the detection of chemical signals found on the female's abdomen (Amrein and Thorne, 2005; Joseph and Carlson, 2015). The male will then proceed to 'sing' a species-specific courtship song by vibrating his wings at a precise frequency (Ritchie *et al.*, 1999). The female assesses the male's song using auditory cues, and recent work suggests that tactile cues in the form of substrate vibrations generated by the male thumping his abdomen on the substrate are also involved at this stage (Fabre *et al.,* 2012). If the female accepts the male's song, he will move behind her and lick her genitals, once again sampling her chemical profile (Joseph and Carlson, 2015). Finally, he will attempt courtship by curling his abdomen towards the female's genitals. If at any point during courtship either fly determines that their potential partner is not an appropriate mate, courtship will be terminated and copulation will not occur (reviewed in: Spieth, 1974). Females may display a variety of rejection behaviours such as extrusion of the ovipositor, spreading her wings,

kicking, or running away from the male (Connolly and Cook, 1973; reviewed in: Spieth, 1974).

1.8 Chemical signaling in *Drosophila* courtship

Drosophila species utilize a variety of chemical cues in addition to CHCs to determine whether or not to court or mate with a potential partner. For example, *cis*-vaccenyl acetate (cVA) is produced in the male ejaculatory bulb and is transferred to the female during mating. It is a volatile compound that can be sensed at range using olfactory receptors by male and female flies alike. *cis*-Vaccenyl acetate is strongly repellent to males, and if a male perceives cVA on a female he will avoid courting her as the presence of cVA serves as an indicator that she has mated recently. *cis*-Vaccenyl acetate will remain on the female for several days following mating, and will diminish as she becomes receptive to mating once more (Jallon, 1984; Datta *et al.,* 2008; Yamamoto and Kohanezawa, 2013).

As in all other insects, *Drosophila* species produce a unique blend of CHCs that are found on their cuticle. The majority of the CHCs produced by *Drosophila* have not been found to have any role as infochemicals, acting mainly to confer the insect with desiccation resistance (reviewed in: Bontonou and Wicker-Thomas, 2014). These compounds are non-covalently bound to the cuticle of the insect and for the most part are sensed only at close range or when directly contacted using gustatory or olfactory receptors (Howard and Blomquist, 2005; reviewed in: Bontonou and Wicker-Thomas, 2014; Joseph and Carlson, 2015). The type and relative abundance of CHCs is dependent on the sex, species, maturity level and overall health of the fly (Howard and Blomquist, 2005; reviewed in: Bontonou and Wicker-Thomas, 2014). The compounds that are active as sex pheromones may act as attractants or as repellent molecules, depending on the identity of the fly that perceives them. Both the type and the ratios of different compounds are important in determining whether a particular CHC profile will be attractive or repulsive to a con- or heterospecific partner (Savarit *et al.*, 1999).

In some *Drosophila* species, flies are sexually monomorphic in the type and abundance of CHCs that they produce while in other species, such as *D. melanogaster*,

flies are sexually dimorphic both in the quantity and type of CHCs that they produce. For example, *D. melanogaster* males produce high levels of 7-tricosene (7-T) and do not produce 7,11-heptacosadiene (7,11-HD) at all, whereas females produce lower levels of 7-T and 7,11-HD represents their most abundant CHC (Antony and Jallon, 1982).

Despite the fact that certain sex pheromones have been shown to induce courtship in *Drosophila* (Billeter *et al.,* 2009), recent evidence indicates that an important role for CHCs may be to act as repellent signals and serve to prevent interspecific mating, rather than as attractants that increase intraspecific courtship and copulation (Dweck *et al.,* 2015). *Drosophila* flies will court potential mates vigorously when they first encounter them (McRoberts and Tompkins, 1987; Griffith and Ejima, 2009), and a negative signal to prevent further courtship is needed once a partner is recognized as a heterospecific. This is supported by the fact that flies whose oenocytes have been ablated that are unable to produce any CHCs are universally attractive to heterospecific flies (Dweck *et al.,* 2015). In *D. melanogaster/D. simulans* pairs, *D. simulans* males that would normally court *D. melanogaster* females at only low levels will court oenocyteless *D. melanogaster* females as vigorously as they court *D. simulans* females (Billeter *et al.*, 2009). When the *D. melanogaster*-specific compound 7,11-HD is added to the cuticle of the oenocytless flies, normal *D. simulans* activity is restored and males avoid courting the *D. melanogaster* females, indicating that 7,11-HD is repellent to *D. simulans* males and that species-specific positive cues are not necessarily required for the initiation of courtship in these *Drosophila* species (Savarit *et al.*, 1999).

1.9 CHC perception in *Drosophila*

CHCs are perceived during courtship using both olfactory and gustatory receptors (Amrein and Thorne, 2005; Joseph and Carlson, 2015). Olfactory receptors are located on the third antennal segments and the maxillary palps and are able to detect CHCs at close range, whereas gustatory receptors are located on the tarsi and proboscis of the fly and must make direct contact with the CHCs in order to detect them (reviewed in: Bontonou and Wicker-Thomas, 2014; Hu *et al.*, 2015). The attractive or repellent nature of a particular compound is highly dependent on the identity of the perceiving fly. In *D. melanogaster,* flies are

sexually dimorphic in their perception of CHCs: males find high levels of 7-T to be repellent, as this is more indicative of a male profile, whereas females are attracted to higher levels of 7-T (reviewed in: Bontonou and Wicker-Thomas, 2014). *Drosophila melanogaster* males also have more (~50) taste bristles than females (~37) on their forelegs, indicating that taste may be more important for males than for females during courtship (Amrein and Thorne, 2005). This makes sense in the context of the *Drosophila* courtship ritual, where males repeatedly tap females on the abdomen to sample their chemical profile.

Several odorant and gustatory receptors have been found to be involved in the direct perception of pheromones, and the removal of these receptors or of the neurons expressing them typically results in dramatically reduced courtship propensity or the loss of discrimination ability in some capacity in males. Three gustatory receptors in particular have been found to be key mediators of courtship in males: gustatory receptor 32a (Gr32a), Gr33a and Gr39a (Amrein and Thorne, 2005). Gr32a is the receptor responsible for the male's ability to discriminate conspecific from heterospecific flies on the basis of their CHC profiles (Fan et al., 2013). When Gr33a is removed, either by a loss-of-function mutation or by ablation of Gr33a-expressing neurons, males court other males at much higher frequency, indicating that removal of the Gr33a receptor causes loss of sex discrimination capability in males. In addition, males lacking active Gr33a lose the ability to discriminate between young and old virgin females (typically males prefer young virgin females over old ones) (Hu et al., 2015). Finally, Gr39a (in addition to Gr33a) has been found to be another gustatory receptor important in the males' ability to distinguish male from female partners (Watanabe et al., 2011).

Several odorant receptors have also been found to be important in the male perception of sex pheromones in *Drosophila*. Key among these are odorant receptor 47b (Or47b) and Or67d. Both of these odorant receptors are known to express the male isoform of *fruitless*, Fru^{M} (Zhuang *et al.*, 2016). Similar to Gr33a, Or47b was found to be responsible for the male preference of younger vs. older virgin females (Zhuang *et al.*, 2016). In addition, Or47b was recently found to be important in the detection of the compound methyl laurate, which might represent a basal attractant CHC across *Drosophila*

species (Dweck *et al.,* 2015). Or67d is responsible for the detection of cVA in both males and females in *Drosophila* (Kurtovic *et al.,* 2007).

1.10 Genes underlying CHC biosynthesis in *Drosophila*

CHCs in *Drosophila* are synthesized using modified fatty acid biosynthetic pathways. Long-chain fatty acids (LCFAs) are used as precursors to modification either through the addition of methyl groups or via desaturation reactions, and these LCFAs can be produced by fatty acid metabolism in either the fat body or the oenocytes (Wicker-Thomas *et al.,* 2015). LCFAs are produced using acetyl-CoA as a precursor, which is then converted to malonyl-CoA using an acetyl-CoA carboxylase (ACC). A fatty acid synthase (FASN) then incorporates further malonyl-CoA subunits onto the acyl-CoA, adding two carbons at a time to the growing chain (Pennanec'h *et al.* 1991; Dembeck *et al.,* 2015). In *Drosophila*, this process results in the production of LCFAs that are 14, 16, or 18 carbons in length. At this point or following transport to the oenocytes, the linear *n*-alkanes may be modified via desaturation or addition of a methyl group (reviewed in: Blomquist and Bagnères, 2010; Dembeck *et al.,* 2015). Oenocyte-specific enzymes then incorporate further units of malonyl-CoA onto the chain in order to form very long-chain fatty acids (VLCFAs). The final step in CHC biosynthesis involves the release of the VLCFA from its CoA side-group, a decarboxylation process that is catalyzed by cytochrome p450 enzymes and which results in the shortening of the chain by one carbon (Dembeck *et al.,* 2015) (see Figure 1.1 for an overview of CHC biosynthesis). The resultant hydrocarbon is then exported to the cuticle via a mechanism that is as yet unknown.

There are several enzymes known to be responsible for modifying linear *n*-alkane chains to produce the known sex pheromones in *Drosophila*. In *D. melanogaster*, the two most abundant male sex pheromones are 7-T and 7-pentacosene (7-P), both of which are alkenes with the double bond at the 7^{th} carbon (Jallon, 1984; reviewed in: Bontonou and Wicker-Thomas, 2014). In females, the most abundant sex pheromones are 7,11-HD and 7,11-nonacosadiene (7,11-ND), both of which are dienes with double bonds at positions 7 and 11 (Jallon, 1984; Pechine *et al.,* 1985). All of these compounds share a double bond at the ω7 position, indicative of a common biosynthetic pathway for the initial desaturation step. This desaturation is carried out by the *Desaturase 1* (*Desat1*) gene, which is expressed

in both the oenocytes and fat body and encodes a Δ^9 fatty acyl-CoA desaturase (Dallerac *et al.,* 2000); the Δ symbol indicates that the desaturase generates a double bond at a fixed position in relation to the carboxyl group of a fatty acid). This desaturase introduces a double bond at the ω7 position of 16-carbon LCFAs and results in the production of monoenes, which are further elongated to form the VLCFAs that are the mature CHCs 7-T and 7-P (reviewed in: Bontonou and Wicker-Thomas, 2014). Some of these monoenes will instead have a second double bond introduced prior to elongation and will become the female compounds 7,11-HD or 7,11-ND (Gleason *et al.,* 2009). Interestingly, *Desat1* expression has been found to regulate not only the production of sex pheromones within *Drosophila* but also how these CHCs are perceived, indicating that a single gene can be responsible for both phenotype and behavioural response to that phenotype (Houot *et al.,* 2010).

Populations of *D. melanogaster* that inhabit African and Caribbean areas produce high amounts of the compound 5,9-heptacosadiene (5,9-HD), which is a positional isomer of 7,11-HD. Populations from all other areas produce low amounts of 5,9-HD and high amounts of 7,11-HD, although these differences have not been found to be involved in the behavioural isolation between these populations (Coyne *et al.,* 1999). *Desat2*, similarly to *Desat1*, encodes a Δ^9 fatty acyl-CoA desaturase, and differences in the expression of this gene between populations of *D. melanogaster* were found to be

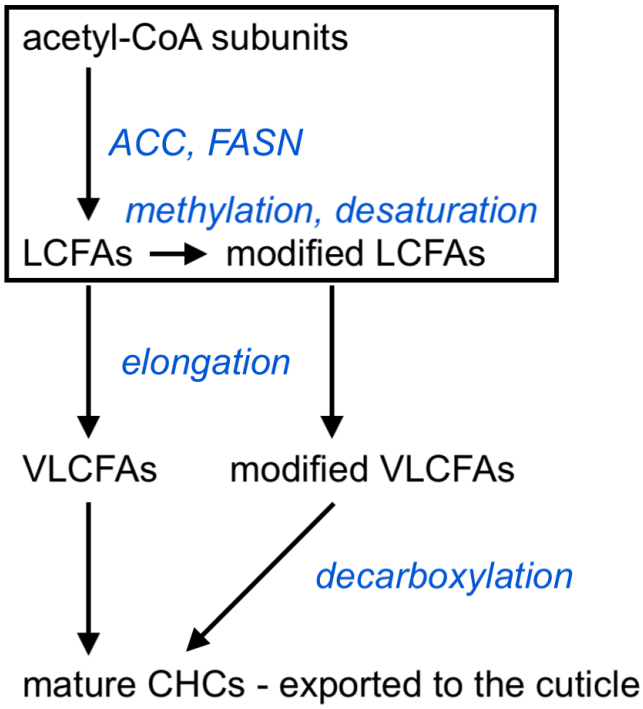


Figure 1.1: Overview of fatty acid biosynthetic processes in *Drosophila*. All processes shown may take place in the oenocytes, while the boxed steps may take place in either the oenocytes or fat body. LCFA = long-chain fatty acid; VLCFA = very long-chain fatty acid; CHC = cuticular hydrocarbon.

responsible for this difference, where females of African and Caribbean strains express *Desat2* but males and all other *D. melanogaster* populations do not (Dallerac *et al.,* 2000; Takahashi *et al.,* 2001).

DesatF, also known as *Fad2,* is responsible for the production of the diene CHCs in *D. melanogaster*. Despite being present in every species in the *melanogaster* subgroup, this gene is only expressed by *D. melanogaster* females (Chertemps *et al.,* 2007). The *DesatF* gene encodes an enzyme that converts ω7 monoene LCFA precursors into dienes by introducing a second double bond at the 11^{th} carbon (ω indicates that the double bond is between the seventh and eighth carbons from the methyl end). The resultant LCFAs are then elongated and decarboxylated into the VLCFAs 7,11-HD and 7,11-ND to become the mature female sex pheromones (Legendre *et al.,* 2008). Chertemps *et al.* discovered that this elongation is carried out by the product of the gene *eloF*, which encodes an elongase that participates in an enzyme complex that is capable of elongating both dienes and monoenes (Chertemps *et al.,* 2007). Finally, a decarbonylase identified as *CYP4G1* is responsible for carrying out the last step in CHC biosynthesis, in which aldehydes that have been released from CoA by an acyl-CoA reductase are converted to mature hydrocarbons (Qiu *et al.,* 2012). This process results in the shortening of the carbon chain by one carbon, which is released as CO_2 (Qiu *et al.,* 2012).

The relative abundance of the CHCs produced by flies is dependent upon the activities of the key enzymes discussed here. If *Desat1* is partially inhibited, relatively lower levels of unsaturated CHCs will be produced (Labeur *et al.,* 2002). Likewise, if *DesatF* is partially inhibited, there will be a buildup of the precursor monounsaturated fatty acids, leading to relatively higher levels of the monoene sex pheromones 7-T and 7-P and lowered levels of the dienes 7,11-HD and 7,11-ND (Wicker-Thomas *et al.,* 2009). In species that do not produce the enzymes necessary to generate dienes, proportionately higher levels of monoenes are produced, as none of the precursors are funneled into the production of dienes.

1.11 Role of generalized metabolism in CHC biosynthesis

Sex pheromone biosynthesis is intricately connected with the lipid metabolism that is responsible for overall fly health (Kuo *et al.,* 2012a; Wicker-Thomas *et al.,* 2015). If fatty acid metabolism in the fat body is hindered, pheromone production is dramatically reduced (Wicker-Thomas *et al.,* 2015), despite the fact that the oenocytes contain all enzymes necessary to synthesize the sex pheromones *de novo* – it seems that crosstalk and LCFA shuttling between the oenocytes and fat body must be required for proper pheromone biosynthesis. Similarly, when juveniles are fed diets containing excess lipid content, adult pheromone levels are seen to drop dramatically (Wicker-Thomas *et al.,* 2015). This likely results from overall LCFA biosynthesis being downregulated in adult flies that receive their necessary lipids from their diet rather than from *de novo* biosynthesis. Since the desaturation enzymes involved in sex pheromone biosynthesis reactions require these LCFAs as precursors in order to make the VLCFAs that will become the mature sex pheromones, adult flies that do not synthesize their own LCFAs are unable to produce high levels of sex pheromones.

Since CHC production utilizes existing fatty acid biosynthetic processes that are necessary for overall health and energy storage of the fly, it has been suggested that, in addition to their roles as species-specific signals, CHCs may represent honest signals of the overall health of the fly (Kuo *et al.,* 2012b). The CHC profile of a fly could therefore assist in determining not only that a potential partner is of the appropriate species, sex and maturity level, but also that it is a high quality individual.

Within-species, CHC profiles of flies can change dramatically based on differences in diet, temperature, and relative humidity (Rouault et al., 2004; Etges *et al.,* 2006; Carvalho *et al.,* 2012; Fedina *et al.,* 2012; Kuo *et al.,* 2012a; Bontonou *et al.,* 2013). Flies reared on different diets begin to mate assortatively after relatively few generations (Dodd, 1989), and this effect was recently found to be mediated by their gut microbial content (Sharon *et al.,* 2010; Najarro *et al.,* 2015). It is possible that the way that flies sense individuals that were raised on the same diet is via some form of chemical communication, implicating differences in the CHC profiles between these flies as a potential candidate.

The adult CHC profile of a fly is partially dependent upon the larval diet (Etges *et al.,* 2006; Wicker-Thomas *et al.,* 2015); since larvae are restricted to the food source that their eggs were laid on, this means that the location a female chooses for laying her eggs can have an effect on the CHC profile, and hence attractiveness, of her adult offspring. If different microbial species are present on different food sources or if different food sources allow for the differential maintenance of the same microbial species, this could lead to assortative mating based solely on where a female lays her eggs and not on underlying genetic factors. This effect would fail to result in the start of speciation between populations if females did not also preferentially lay their eggs on the same food source they were raised on, as random oviposition sites would allow for free gene flow between populations, but it offers an intriguing avenue of research into how traits involved in mate recognition can change rapidly in response to environmental cues.

1.12 Factors influencing CHC profiles in *Drosophila*

The range of unique CHCs that can be produced by a fly is dependent upon its genetic background, while the quantity of specific compounds can also be modified via interaction with the environment. Since the CHC profile of a fly is key in determining its reproductive success, both the genes underlying CHC production and how this production can change in response to different environmental factors are important in studying the evolution of mate choice and behavioural isolation in *Drosophila*. In examining the evolution of behavioural prezygotic barriers, it is therefore important to look at both the existing genetic factors that are present between species that cause them to be behaviourally isolated, and to examine how potential barriers may come to evolve in the first place. In this study, I therefore aim to: a) identify genes that underlie species differences in CHC production between two reproductively isolated sibling species of *Drosophila*: *D. melanogaster* and *D. simulans*, and b) look at how the CHC profiles within *D. melanogaster* change in response to differences in both diet and microbial content. This should yield insight into the genetics underlying extant behavioural isolation, as well as probe into how phenotypic plasticity in sexually selected traits may allow for the onset of behavioural isolation within a single species group.

1.13 Bibliography

Ali, M.F. and Morgan, E.D. **1990**. Chemical communication in insect communities: A guide to insect pheromones with special emphasis on social insects. *Biological Reviews***. 65**: 227-247.

Amrein, H. and Thorne, N. **2005**. Gustatory perception and behavior in *Drosophila melanogaster*. *Current Biology*. **15**:673-684.

Andersson, M., Simmons, L.W. **2006**. Sexual selection and mate choice. *Trends in Ecology and Evolution*. **21**:296-302.

Antony, C. and Jallon, J-M. **1982**. The chemical basis for sex recognition in *Drosophila melanogaster*. *Journal of Insect Physiology*. **28**:873-880.

Billeter J.C., Atallah J., Krupp J.J., Millar J.G. and Levine J.D. **2009**. Specialized cells tag sexual and species identity in *Drosophila melanogaster*. *Nature*. **461**:987–991.

Blomquist, G.J. and Bagnères, A-G. **2010**. Insect hydrocarbons: Biology, biochemistry, and chemical ecology. Cambridge: Cambridge University Press.

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*. **26**:2008-2020.

Bontonou, G. and Wicker-Thomas, C. **2014**. Sexual communication in the *Drosophila* genus. *Insects*. **5**:439-458.

Burley, N. **1977**. Parental investment, mate choice, and mate quality. *Proceedings of the National Academy of Sciences USA*. **74**:3476-3479.

Carvalho, M., Sampaio, J.L., Palm, W., Brankatschk, M., Eaton, S. and Shevchenk, A. **2012**. Effects of diet and development on the *Drosophila* lipidome. *Molecular Systems Biology*. **8**:600.

Cattani, M.V and Presgraves, D.C. **2012**. Incompatibility between X chromosome factor and pericentric heterochromatic region causes lethality in hybrids between *Drosophila melanogaster* and its sibling species. *Genetics*. **191**: 549-559.

Chertemps, T., Duportets, L., Labeur, C., Ueda, R., Takahashi, K., Saigo, K. and Wicker-Thomas, C. **2007**. A female-biased expressed elongase involved in long-chain hydrocarbon biosynthesis and courtship behavior in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA*. **104**:4273-4278.

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*. **37**:822-830.

Connolly, K. and Cook, R. **1973**. Rejection responses by female *Drosophila melanogaster*: their ontogeny, causality and effects upon the behaviour of the courting male. *Behaviour*. **44**:142-166.

Coyne, J.A., Wicker-Thomas, C. and Jallon, J-M. **1999**. A gene responsible for a cuticular hydrocarbon polymorphism in *Drosophila melanogaster*. *Genetics Research*. **73**:189–203.

Coyne, J.A. and Orr, H.A. **2004**. Speciation. Sunderland, Mass: Sinauer Associates.

Craig, J.F. **1980**. Behaviour and evolution in the genus *Euplectes*. *Journal fur Ornithologie*. **121**:144-161.

Curtis, H. and Barnes, N.S. **1989**. Biology, 5ed. New York: W.H. Freeman.

Darwin, C. **1859**. On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. London: J. Murray.

Datta, S.R., Vasconcelos, M.L., Ruta, V., Luo, S., Wong, A., Demir, E., Flores, J., Balonze, K., Dickson, B.J. and Axel, R. **2008**. The *Drosophila* pheromone cVA activates a sexually dimorphic neural circuit. *Nature*. **452**:473-477.

Dallerac, R., Labeur, C., Jallon, J.M., Knipple, D.C., Roelofs, W.L. and Wicker-Thomas, C. **2000**. A delta 9 desaturase gene with a different substrate specificity is responsible for the cuticular diene hydrocarbon polymorphism in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA*. **97**:9449-9454.

Dembeck, L.M., Böröczky, K., Huang, W., Schal, C., Anholt, R.R. and Mackay, T.F. **2015**. Genetic architecture of natural variation in cuticular hydrocarbon composition in *Drosophila melanogaster*. *eLife*. **4**:e09861.

Dettner, K. and Liepert, C. **1994**. Chemical mimicry and camouflage. *Annual Review of Entomology*. **39**:129-154.

Dobzhansky, T. **1937**. Genetics and the Origin of Species. New York: Columbia Univ. Press.

Dodd, D.M. **1989**. Reproductive isolation as a consequence of adaptive divergence in *Drosophila pseudoobscura*. *Evolution*. **43**:1308-1311.

Dweck, H.K., Ebrahim, S.A., Thoma, M., Mohamed, A.A., Keesey, I.W., Trona, F., Lavista-Llanos, S., Svatoš, A., Sachse, S., Knaden, M. and Hansson, B.S. **2015**. Pheromones mediating copulation and attractiveness in *Drosophila*. *Proceedings of the National Academy of Sciences USA*. **112**:2829-2835.

Ellis, L.L. and Carney, G.E. **2009**. *Drosophila melanogaster* males respond differently at the behavioural and genome-wide levels to *Drosophila melanogaster* and *Drosophila simulans* females. *Journal of Evolutionary Biology*. **22**:2183-2191.

Etges, W.J., Veenstra, C.L. and Jackson, L.J. **2006**. Premating isolation is determined by larval rearing substrates in cactophilic *Drosophila mojavensis*. VII. Effects of larval dietary fatty acids on adult epicuticular hydrocarbons. *Journal of Chemical Ecology*. **32**:2629-2646.

Fabre, C.C., Hedwig, B., Conduit, G., Lawrence, P.A., Goodwin, S.F. and Casal, J. **2012**. Substrate-borne vibratory communication during courtship in *Drosophila melanogaster*. *Current Biology*. **22**:2180-2185.

Fan, P., Manoli, D.S., Ahmed, O.M., Chen, Y., Agarwal, N., Kwong, S., Cai, A.G., Neitz, J., Renslo, A., Baker, B.S. and Shah, N.M. **2013**. Genetic and neural mechanisms that inhibit *Drosophila* from mating with other species. *Cell*. **154**:89–102.

Fedina, T.Y., Kuo, T-H., Dreisewerd, K., Dierick, H.A., Yew, J.Y. and Pletcher, S.D. **2012**. Dietary effects on cuticular hydrocarbons and sexual attractiveness in *Drosophila*. *Public Library of Science*. **7**:e49799.

Friesen, V.L., Smith, A.L., Gomez-Diaz, E., Bolton, M., Furness, R.W., Gonzalez-Solis, J. and Monteiro, L.R. **2007**. Sympatric speciation by allochrony in a seabird. *Proceedings of the National Academy of Sciences USA*. **104**:18589-18594.

Ginzel, M.D. **2010**. Encyclopedia of Animal Behavior. Olfactory signals. Academic Press: Cambridge, Massachusetts. 584-588.

Gleason, J.M., James, R.A., Wicker-Thomas, C. and Ritchie, M.G. **2009**. Identification of quantitative trait loci function through analysis of cuticular hydrocarbons differing between *Drosophila simulans* and *Drosophila sechellia* females. *Heredity*. **103**:416-424.

Greene, M.J. and Gordon, D.M. **2003**. Social insects: Cuticular hydrocarbons inform task decisions. *Nature*. **423**:32.

Greenspan, R.J. **1995**. Understanding the genetic construction of behavior. *Scientific American*. **272**:74–79.

Griffith, L.C. and Ejima, A. **2009**. Multimodal sensory integration of courtship stimulating cues in *Drosophila melanogaster*. *Annals of the New York Academy of Sciences*. **1170**:394-398.

Horth, L. **2007**. Sensory genes and mate choice: Evidence that duplications, mutations, and adaptive evolution alter variation in mating cue genes and their receptors. *Genomics*. **90**:159-175.

Houot, B., Bousquet, F. and Ferveur, J.F. **2010**. The consequences of regulation of *desat1* expression for pheromone emission and detection in *Drosophila melanogaster*. *Genetics*. **185**:1297-1309.

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

Hu, Y., Han, Y., Shao, Y., Wang, X., Ma, Y., Ling, E. and Xue, L. **2015**. *Gr33a* modulates *Drosophila* male courtship preference. Scientific Reports. **5**: 7777.

Jackson, B.D. and Morgan, E.D. **1993**. Insect chemical communication: Pheromones and exocrine glands of ants. *Chemoecology*. **4**:125-144.

Jallon, J. M. **1984**. A few chemical words exchanged by *Drosophila* during courtship and mating. *Behavioral Genetics*. **14**:441–478.

Johansson, B.G. and Jones, T.M. **2007**. The role of chemical communication in mate choice. *Biological Reviews of the Cambridge Philosophical Society*. **82**:265-289.

Joseph, R.M. and Carlson, J.R. **2015**. *Drosophila* chemoreceptors: A molecular interface between the chemical world and the brain. *Trends in Genetics*. **31**:683-695.

Ko, W-Y., David, R.M. and Akashi, H. **2003**. Molecular phylogeny of the *Drosophila melanogaster* species subgroup. *Journal of Molecular Evolution*. **57**:562-573.

Kokko, H. and Jennions, M.D. **2014**. The relationship between sexual selection and sexual conflict. *CSH Perspectives in Biology*. **6**:a017517.

Kozak, G.M., Reisland, M. and Boughmann, J.W. **2008**. Sex differences in mate recognition and conspecific preference in species with mutual mate choice. *The Society for the Study of Evolution*. **63**:353-365.

Kuo, T-H., Fedina, T.Y., Hansen, I., Dreisewerd, K., Dierick, H.A., Yew, J.Y. and Pletcher, S.D. **2012a**. Insulin signaling mediates sexual attractiveness in *Drosophila*. *Public Library of Science*. **8**:e1002684.

Kuo, T-H., Yew, J.Y., Fedina, T.Y., Dreisewerd, K., Dierick, H.A. and Pletcher, S.D. **2012b**. Aging modulates cuticular hydrocarbons and sexual attractiveness in *Drosophila melanogaster*. *Journal of Experimental Biology*. **215**:814-821.

Kurtovic, A., Widmer, A. and Dickson, B.J. **2007**. A single class of olfactory neurons mediates behavioural responses to a *Drosophila* sex pheromone. *Nature*. **446**:542-546.

Labeur, C., Dallerac, R. and Wicker-Thomas, C. **2002**. Involvement of the *desat1* gene in the control of *Drosophila melanogaster* pheromone biosynthesis. *Genetica*. **114**:269-274.

Leclerc, G-L., Compte de Buffon. **1766**. Histoire Naturelle, Générale et Particulière, avec la description du Cabinet du Roi. Paris: Imprimerie Royale.

Legendre, A., Miao, X.-X., Da Lage, J.-L. and Wicker-Thomas, C. **2008**. Evolution of a desaturase involved in female pheromonal cuticular hydrocarbon biosynthesis and courtship behavior in *Drosophila*. *Insect Biochemistry and Molecular Biology*. **38**:244-255.

Liepert, C. and Dettner, K. **1996**. Role of cuticular hydrocarbons of aphid parasitoids in their relationship to aphid-attending ants. *Journal of Chemical Ecology*. **22**:695-707.

Ludlow, A.M. and Magurran, A.E. **2006**. Gametic isolation in guppies (*Poecilia reticulata*). *Proceedings of the Royal Society B: Biological Sciences*. **273**:2477-2482.

Najarro, M.A., Sumethasorn, M., Lamoureux, A. and Turner, T.L. **2015**. Choosing mates based on the diet of your ancestors: replication of non-genetic assortative mating in *Drosophila melanogaster*. *PeerJ*. **3**:e1173.

Martin, O.Y., and Hosken, D.J. **2003**. The evolution of reproductive isolation through sexual conflict. *Nature*. **423**:979-982.

Matute, D.R. **2014**. The magnitude of behavioral isolation is affected by characteristics of the mating community. *Ecology and Evolution*. **4**:2945-2956.

Matute, D.R. and Coyne, J.A. **2010**. Intrinsic reproductive isolation between two sister species of *Drosophila*. *Evolution*. **64**:903–920

Mayr, E. **1942**. Systematics and the Origin of Species. New York: Columbia Univ. Press.

McNabney, D.R. **2012**. The genetic basis of behavioral isolation between *Drosophila mauritania* and *D. sechellia*. *Evolution*. **66**:2182-2190.

McRobert, S.P. and Tompkins, L. **1987**. The effect of light on the sexual behavior of *Drosophila affinis*. *Behavioral and Neural Biology*. **47**:151-157.

Moehring, A.J., Llopart, A., Elwyn, S., Coyne, J.A. and Mackay, T.F. **2006**. The genetic basis of postzygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to hybrid male sterility. *Genetics*. **173**:225–233

Oritz-Barrientos, D., Counterman, B.A. and Noor, M.A. **2004**. The genetics of speciation by reinforcement. *Public Library of Science*. **2**:e416.

Orr, M.R. and Smith, T.B. **1998**. Ecology and speciation. *Trends in Ecology and Evolution*. **13**:502-506.

Pechine, J.M., Perez, F., Antony, C. and Jallon, J.M. **1985**. A further characterization of *Drosophila* cuticular monoenes using a mass spectrometry method to localize double bonds in complex mixtures. *Analytical Biochemistry*. **145**:177-182.

Pener, M.P. and Yerushalmi, Y. **1998**. The physiology of locust phase polymorphism: an update. *Journal of Insect Physiology*. **44**:365-377.

Pennanec'h, M., Ferveur, J-F., Pho D-B. and Jallon J-M. **1991**. Insect fatty acid related pheromones: A review of their biosynthesis, hormonal regulation and genetic control. *Annales de la Societe Entomologique de France*. **27**:245-263.

Qiu, Y., Tittiger, C., Wicker-Thomas, C., le Goff, G., Young, S., Wajnberg, E., Fricaux, T., Taquet, N., Blomquist, G.J. and Feyereisen, R. **2012**. An insect-specific P450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis. *Proceedings of the National Academy of Sciences USA*. **109**:14858-14863.

Rhymer, J.M. and Simberloff, D. **1996**. Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*. **27**:83– 109.

Rieseberg, L.H., Widmer, A., Arntz, A.M. and Burke, J.M. **2002**. Directional selection is the primary cause of phenotypic diversification. *Proceedings of the National Academy of Sciences USA*. **99**:12242-12245.

Ritchie, M.G., Halsey, E.J. and Gleason, J.M. **1999**. *Drosophila* song as a species-specific mating signal and the behavioural importance of Kyriacou & Hall cycles in *D. melanogaster* song. *Animal Behaviour*. **58**:649-657.

Rouault, J.D., Marican, C., Wicker-Thomas, C. and Jallon, J-M. **2004**. Relations between cuticular hydrocarbon (HC) polymorphism, resistance to desiccation and breeding temperature; a model for HC evolution in *D. melanogaster* and *D. simulans*. *Genetica*. **120**:195-212.

Savarit F., Sureau G., Cobb M. and Ferveur J.F. **1999**. Genetic elimination of known pheromones reveals the fundamental chemical bases of mating and isolation in *Drosophila*. *Proceedings of the National Academy of Sciences USA*. **96**:9015–9020.

Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A. **2014**. Genomics and the origin of species. *Nature Reviews Genetics*. **15**:176-192.

Servedio, M.R. **2016**. Geography, assortative mating, and the effects of sexual selection on speciation with gene flow. *Evolutionary Applications*. **9**:91-102.

Sharon, G., Segal, D., Ringo, J.M., Hefetz, A., Zilber-Rosenberg, I. and Rosenberg, E. **2010**. Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA*. **107**:20051-20056.

Sledge, M.F., Boscaro, F. and Turillazzi, S. **2001**. Cuticular hydrocarbons and reproductive status in the social wasp *Polistes dominulus*. *Behavioural Ecology and Sociobiology*. **49**:401-409.

Sobel, J.M., Chen, G.F., Watt, L.R. and Schemske, D.W. **2010**. The biology of speciation. *Evolution*. **64**:295-315.

Spieth, H.T. **1974**. Courtship behaviour in *Drosophila*. *Annual Review of Entomology*. **19**:185-405.

Starck, J.M. **1991**. Biogeography and life history of *Turnix suscitator*. Small adult body size as a consequence of selection for rapid growth. *Journal of Zoological Systematics and Evolutionary Research*. **29**:213-237.

Takahashi, A., Tsaur, S.C., Coyne, J.A. and Wu, C.I. **2001**. The nucleotide changes governing cuticular hydrocarbon variation and their evolution in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA*. **98**:3920–3925.

Trivers, R.L. **1972**. Parental investment and sexual selection. In Sexual Selection & the Descent of Man. New York: Aldine de Gruyter. 136-179.

Yamamoto, D. and Kohanezawa, M. **2013**. Genes and circuits of courtship behaviour in *Drosophila* males. *Nature Reviews: Neuroscience*. **14**:681-692.

Ward, J.L. and McLennan, D.A. **2009**. Mate choice based on complex visual signals in the brook stickleback, *Culaea inconstans*. *Behavioural Ecology*. **20**: 1323-1333.

Watanabe, K., Toba, G., Koganezawa, M. and Yamamoto, D. **2011**. Gr39a, a highly diversified gustatory receptor in *Drosophila*, has a role in sexual behavior. *Behavioral Genetics*. **41**:746–753.

Wicker-Thomas, C., Guenachi, I. and Keita, Y.F. **2009**. Contribution of oenocytes and pheromones to courtship behaviour in *Drosophila*. *BMC Biochemistry*. **10**:21.

Wicker-Thomas, C., Garrido, D., Bontonou, G., Napal, L., Mazuras, N., Denis, B., Rubin, T., Parvy, J-P. and Montagne, J. **2015**. Flexible origin of hydrocarbon/ pheromone precursors in *Drosophila melanogaster*. *Journal of Lipid Research*. **56**:2094-2101.

Zhuang, L., Sun, Y., Hu, M., Chenxi, W., La, X., Chen, X., Feng, Y., Wang, X., Hu, Y. and Xue, L. **2016**. Or47b plays a role in *Drosophila* males' preference for younger mates. *Public Library of Science*. **6**:160086.

Chapter 2

2 Genes underlying species differences in CHC production between *Drosophila melanogaster* and *D. simulans*

2.1 Introduction

How species diverge and remain distinct is of key import in the study of the evolutionary processes that give rise to these groups. Of particular interest are the genetic mechanisms underlying the isolating barriers that arise between different groups, especially those that exist between groups that have diverged relatively recently. The genetic differences found between more recently formed species are more likely to be represent some of the most basal possible changes that may give rise to or maintain these species as distinct groups. Currently, little is known about the genetic basis of behavioural isolation between species and the genes that may be involved early in the speciation process that result in changes in behaviour and mating propensity. In *Drosophila*, courtship utilizes chemical communication at several steps (reviewed in: Fernández and Kravitz, 2013), indicating that the differential expression or perception of chemical cues may represent a key factor in the behavioural phenotype expressed by different *Drosophila* species that results in their reproductive isolation. Here, I sought to discover the genetic basis for the isolating mating behaviour observed between *Drosophila melanogaster* and *D. simulans* by exploring the genes involved in differential sex pheromone biosynthesis between these species.

2.1.1 Reproductive isolation between *D. melanogaster* and *D. simulans*

Drosophila melanogaster and *D. simulans* are a well-suited model system for exploring the genetic basis of chemical communication and behavioural isolation. This species pair diverged in sympatry approximately 5.4 mya (Sousa-Neves and Rosas, 2010). As in other species in the *melanogaster* subgroup, they are both prezygotically and postzygotically isolated from one another. Sterile hybrids that are the same sex as the *D. melanogaster* parent can be produced, while hybrids of the sex opposite to the *D. melanogaster* parent

are inviable and fail to develop (Yamamoto, 1992; Matute *et al.,* 2014). Males from both species will court heterospecific females, but at reduced frequency compared to conspecific courting (Ellis and Carney, 2009). However, the species pair is asymetrically behaviourally isolated due to female behaviour: *D. melanogaster* females will accept *D. simulans* males at low frequency, but the reciprocal cross is observed only very rarely (Cobb and Jallon, 1990). This behavioural isolation is due, at least in part, to differences in sex pheromones between the species (Savarit *et al.,* 1999; Coyne, 1996).

2.1.2 Differences in chemical cues between *D. melanogaster* and *D. simulans*

Drosophila species, as in many other insects, utilize cuticular hydrocarbons (CHCs) both for desiccation resistance and to communicate during courtship (reviewed in: Blomquist and Bagnères, 2010; Howard and Blomquist, 2005). These fatty-acid derivative molecules are loosely associated with the cuticle of the fly, and are typically detected using olfactory and gustatory receptors only when flies come into close contact (Joseph and Carlson, 2015; Amrein and Thorne, 2005). *Drosophila melanogaster* is sexually dimorphic in the types of CHCs that it produces, whereas *D. simulans* is sexually monomorphic. Typically, males from monomorphic *Drosophila* species are reluctant to court females from dimorphic species (reviewed in: Bontonou and Wicker-Thomas, 2014). Four sex pheromones in particular that are produced by one or both of these species: 7,11-heptacosadiene (7,11-HD), 7,11-nonacosadiene (7,11-ND), 7-tricosene (7-T) and 7-pentacosene (7-P) have been studied extensively. Since *D. melanogaster* and *D. simulans* respond differently to the presence or ratios of these key CHCs, this differential perception can explain part of the behavioural isolation between these species.

7,11-Heptacosadiene is produced solely by *D. melanogaster* females, and represents the most abundant CHC in these flies (Antony and Jallon, 1982). This compound acts as a stimulant to *D. melanogaster* males, but is strongly repellent to *D. simulans* males (Coyne et al., 1994). *Drosophila melanogaster* females lacking 7,11-HD are courted vigorously by *D. simulans* males, indicating that it serves as one of the major species-

specific antiaphrodesiac signals between *D. melanogaster* and *D. simulans* (Billeter *et al.,* 2009).

Similar to 7,11-HD, 7,11-ND is also only produced by *D. melanogaster* females, and it is known to stimulate conspecific males and to act as an antiaphrodesiac to heterospecifics, though to a lesser degree than 7,11-HD (Jallon, 1984). Both 7,11-HD and 7,11-ND were found to induce wing vibration in *D. melanogaster* males, with 7,11-HD causing the most profound stimulatory effect (Antony and Jallon, 1982).

7-Tricosene and 7-P, found on the cuticles of male flies in all species of the *melanogaster* subgroup (reviewed in: Bontonou and Wicker-Thomas, 2014), comprise more than half of total male CHCs. The ratio between these compounds has been found to be dependent upon the latitude at which fly populations live, consistent with the fact that 7-P provides greater resistance to desiccation and is typically found at higher amounts in fly populations that live in warmer environments (Rouault et al., 2004). In *D. melanogaster*, both males and females produce 7-T and 7-P, however females produce very low levels of both compounds. *Drosophila melanogaster* females are stimulated by exposure to high levels of 7-T, whereas 7-T acts as an antiaphrodesiac to *D. melanogaster* males, allowing them to avoid male-male courtship (Thistle *et al.,* 2012). In *D. simulans*, 7-T is the most abundant sex pheromone found on the cuticles of both male and female flies, who differ only in the relative abundances and not the identities of their CHCs (reviewed in: Bontonou and Wicker-Thomas, 2014). Male *D. melanogaster* flies interacting with *D. simulans* females are repelled due to the presence of high amounts of 7-T, indicative of a male profile (Thistle *et al.,* 2012). 7-Tricosene has also been found to stimulate wing vibration in *D. simulans* males (Jallon, 1984), but not in *D. melanogaster* males (Antony and Jallon, 1982).

Hybrid females produced by *D. melanogaster* females that have been crossed to *D. simulans* males have overall CHC profiles that are intermediate between the parental profiles, but that are semi-dominant for the *D. melanogaster* profile (Coyne, 1996; Pardy, 2012). These hybrid flies produce high levels of 7,11-HD and low amounts of 7,T, which is more consistent with a *D. melanogaster* female profile. Other, less well-studied CHCs

also differ in amount in the hybrid vs. parental CHC profiles, which may be important as the way that flies perceive CHC profiles is dependent not only upon the types of compounds found but upon their relative abundances compared to other CHCs and the overall bouquet of compounds that are present (Ferveur and Sureau, 1996).

2.1.3 Chromosomal regions underlying CHC differences between *D. melanogaster* and *D. simulans*

Although some key genes have been identified (see section 1.1.0), the genetic basis of CHC production is in general not well known. Previous work has identified that the genes underlying the CHC profiles of *D. melanogaster/D. simulans* hybrids map entirely to the 3rd chromosome (Coyne, 1996). Deficiency mapping was used to walk along the third chromosome and identify smaller regions that may be responsible for the hybrid CHC profile, uncovering 5 regions that each contain between 63 and 230 genes (Pardy, 2012). These regions must each contain at least one gene that is responsible for some of the species differences in CHC production that is observed between *D. melanogaster* and *D. simulans*. These maps are important not only in identifying the genes that may be acting between species, but also in looking at those genes that underlie CHC production in general. Since the behavioural isolation between *D. melanogaster* and *D. simulans* is partially modulated by their different CHC profiles, it is necessary to discover which specific genes within the large regions uncovered by deficiency mapping are contributing to the differences in CHC profiles between these species.

2.1.4 Candidate gene disruptions

In order to discover which individual genes may be acting to affect species differences in CHC production between *D. melanogaster* and *D. simulans*, a version of deficiency mapping was used in which individual candidate genes (rather than large regions) are disrupted, as per Pasyukova *et al*., 2000. *Drosophila melanogaster* fly stocks are available in which individual genes have been rendered nonfunctional, either by the insertion of a transposable element or deletion of nucleotides. These stocks are hemizygous (have only one homolog) for the candidate gene of interest, which is maintained over a balancer

chromosome to prevent the recovery of recombinant offspring (the balancer chromosome itself reduces recombination rates, and if recombination does occur recombinant offspring are not viable). When *D. melanogaster* stocks containing a gene disruption are crossed to wild-type (WT) *D. simulans* flies, hybrids are produced that differ from WT hybrids only in the absence of the *D. melanogaster* copy of the candidate gene. If the CHC profile of disrupted hybrids differs significantly from the WT hybrid, this disrupted gene must therefore be responsible for the difference.

In my research, I tested individual candidate genes on the 3rd chromosome to attempt to narrow down which genes contribute to CHC differences between *D. melanogaster* and *D. simulans*. In choosing candidates, I searched within areas uncovered by Pardy (2012), as well as along the entire length of the 3rd chromosome, for candidate genes. Candidate genes were chosen based on (1) whether they are involved in fatty acid biosynthetic processes, namely those putatively involved in desaturation, elongation, and decarboxylation reactions, identified either through experimentation or inferred function based on protein domains; or (2) if they are known to interact with or take part in pathways shared with the products of those genes already known to be involved in CHC biosynthesis, such as *Desat1* (Dallerac *et al.,* 2000), *Desat2* (Coyne *et al.,* 1999) and *DesatF* (*Fad2*) (Chertemps *et al.,* 2007).

2.2 Methods

2.2.1 Maintenance of *Drosophila* stocks

Drosophila stocks were maintained on a standard cornmeal-yeast diet (Bloomington Drosophila Stock Center standard recipe) at 24°C on a 14:10 hour light:dark cycle at 76% relative humidity. Wild-type (WT) *D. melanogaster* line BJS (London, ON) was obtained from Dr. B. Sinclair and WT *D. simulans* line FC (Florida City, USA) was obtained from Dr. J. Coyne. Balancer stock 3703 ($w^{1118}/Dp(1;Y)y^{+}$; CyO/nub^{1} b^{1} sna^{Sco} lt^{1} stw^{3}; $MKRS/TM6B$, Tb^{1}) and gene disruption stocks (see Table 2.1) were obtained from the Bloomington *Drosophila* Stock Center.

2.2.2 *Drosophila* crosses

Newly-emerged flies (0-8h) were collected under light CO_2 anesthesia and separated by sex to ensure virginity. *Drosophila melanogaster* disruption line flies (*Dis*) were crossed with a balancer stock (*Bal*) to generate F1 offspring that have a copy of chromosome 3 containing a disrupted gene and a homologous balancer chromosome that prevents recovery of recombinant offspring (*Dis/Bal*). Offspring were selected on the basis of the phenotype imparted by a dominant visible marker present on the balancer chromosome. *Dis/Bal D. melanogaster* virgin female flies were then crossed with either WT male *D. melanogaster* BJS (3 crosses/line, 5 females x 5 males) or WT male *D. simulans* FC (a minimum of 10 crosses and an average of 15 crosses/line, 10 females x 25 males) flies. A greater number of females and males were used in the interspecies cross due to the reduced mating activity for this cross. Female F1 hybrid offspring were collected under light CO_2 anesthesia and separated 0-8h after eclosion based on the presence (inherited *Bal*) or absence (inherited *Dis*) of the dominant visible marker. See figure 2.1 for an overview of the crossing scheme.

In many cases, crosses between *Dis/Bal D. melanogaster* females and WT *D. simulans* males were unsuccessful. These crosses resulted in hybrids that either failed to eclose or

that eclosed but did not survive to testing age (5 days) in high enough numbers. For some stocks, no larvae were observed in any interspecific cross (see Table 2.1).

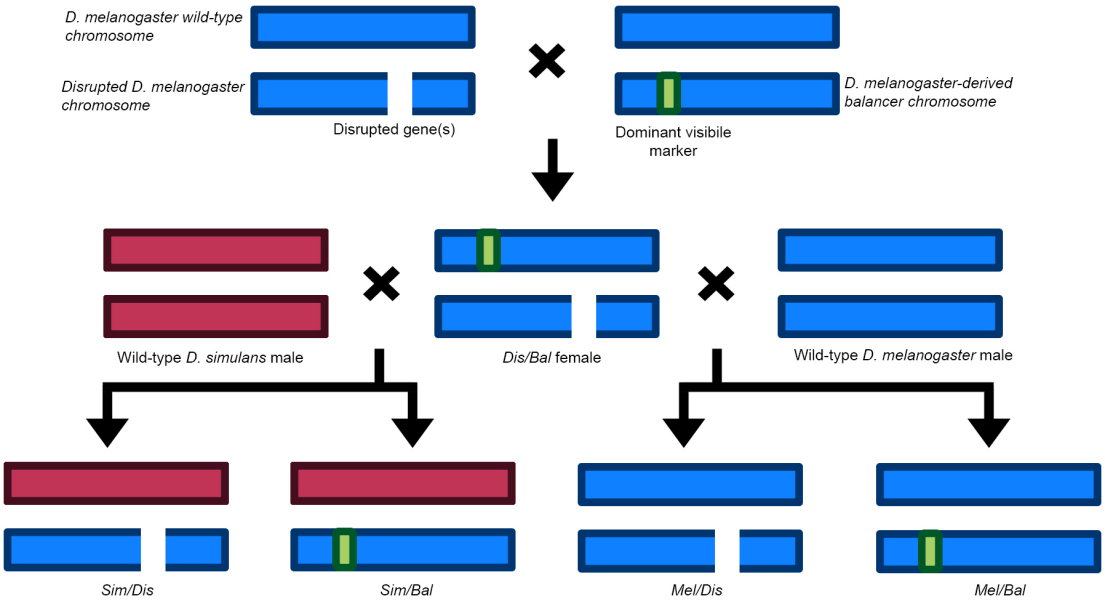


Figure 2.1: Crossing scheme for testing candidate genes involved in CHC biosynthesis. Four possible genotypes, shown at the bottom, are produced by inter- and intraspecific crosses. *Sim/Dis* individuals have no copies of the *D. melanogaster* copy of the candidate gene; *Sim/Bal* has a single copy; *Mel/Dis* has a single copy; and *Mel/Bal* has two copies.

2.2.3 Candidate genes identified using STRINGdb

Candidate genes tested are listed in Table 2.1. As mentioned, candidate genes were chosen from regions spanning the third chromosome that either: a) occur within regions found to significantly affect CHC profiles between *D. melanogaster* and *D. simulans* (Pardy, 2012); b) are involved in fatty acid biosynthetic processes; or c) are known to interact with the protein products of genes that have been previously shown to be involved in species differences in CHC production between *D. melanogaster* and *D. simulans*. The genes *Desat1*, *Desat2* and *DesatF* are known to be involved in CHC biosynthesis in *Drosophila* (Dallerac *et al.,* 2000; Coyne *et al.,* 1999; Chertemps *et al.,* 2007) and were used as starting points to query STRINGdb for any known interacting partners. STRING is a freely available service that accepts a protein as input and searches for other proteins that may associate with the query. STRING searches for evidence of this interaction via experimental evidence and using known homologs of the protein of interest in other species (Jensen *et al.,* 2009), and provides a good starting point in looking for proteins that may interact with the protein of interest. Since there is already strong evidence for the involvement of *Desat1*, *Desat2* and *DesatF* in the synthesis of the *D. melanogaster* sex pheromones, any proteins that interact with them represent good candidates for other proteins that may be involved in the pheromone biosynthetic process. I therefore looked for the top interacting partners of the products of these three genes based on the STRINGdb output that are also located on the third chromosome and used these as candidate genes.

Table 1: Candidate genes for species differences in CHC production between *D. melanogaster* and *D. simulans*

Stock[1]	Genotype[1]	Affected gene[2]	Gene function[2]	Cytological region[2]
1 - Stocks that produced enough interspecific offspring to test[3]				
25101	*w*; st*[1] faf*[BX3]/TM6B, Tb[1]*	*CG1945 - faf*	thiol-dependent ubiquitin-specific protease activity	100D2-100D2
51074	*y[1 w*; Mi{y[+mDint2]=MIC}sit[MI08264]/TM 3, Sb[1] Ser[1]*	*CG5278 - sit*	fatty acid elongase activity	94B4-94B4
54496	*y[1] w*; Mi{y[+mDint2]=MIC}bond[MI09426] CR44062[MI09426]*	*CG6921 - bond*	fatty acid elongase activity	94B4-94B4
		CR44062	uncharacterized	94B4-94B4
51122	*y[1] w*; Mi{y[+mDint2]=MIC}CG33110[MI08 895]/TM3, Sb[1] Ser[1]*	*CG33110*	fatty acid elongase activity	94B4-94B4
16264	*y[1] w[1118]; PBac{y[+mDint]=3HPy[+]}C040*	*CG42857*	ucharacterized	85E10-85E10
35100	*y[1] w*; Mi{y[+mDint2]=MIC}Oct-TyrR[MI01223]/TM3, Sb[1] Ser[1]*	*CG7485 - Oct-TyrR*	Tyramine 1 class receptor	79B1-79B2

Stock[1]	Genotype[1]	Affected gene[2]	Gene function[2]	Cytological region[2]
14660	*y[1]; P{y[+mDint2] w[BR.E.BR]=SUPor-P}verm[KG07819] ry[506]/TM3, Sb[1] Ser[1]*	*CG8756 - verm*	Chitin-binding	76C1-76C3
27232	*y[1] w[*]; P{w[+mC]=EP}SREBP[G8480]*	*CG8522 - SREBP*	Transcriptional activator	76C5-76C6
19706	*y[1] w[67c23]; P{w[+mC] y[+mDint2]=EPgy2}Baldspot[EY05228]*	*CG3971 - Baldspot*	fatty acid elongase activity	73B4-73B5
12612	*w[1118]; P{w[+mGT]=GT1}CG5946[BG01087]*	*CG5946*	oxidoreductase activity	68E1-68E1
2 - Stocks that produced interspecific offspring that died before testing age[3]				
22471	*y[1] w[67c23]; P{w[+mC] y[+mDint2]=EPgy2}E(z)[EY21318]*	*CG6502 - E(z)*	Histone methyltransferase; involved in development	67E5-67E5
33478	*y[1] w[*]; Mi{y[+mDint2]=MIC}Acp76A[MI02019]*	*CG3801 - Acp76A*	Non-inhibitory serpin serine endopeptidase inhibitor family	75F5-75F5
17835	*w[1118]; PBac{w[+mC]=RB}CG10096[e00276] CG10097[e00276]*	*CG10096*	fatty-acyl-CoA reductase (alcohol-forming) activity	87B9-87B9
		CG10097	fatty-acyl-CoA reductase (alcohol-forming) activity	87B9-87B9

Stock[1]	Genotype[1]	Affected gene[2]	Gene function[2]	Cytological region[2]
33155	*y*^{1} *w**; *Mi{y*$^{+mDint2}$*=MIC}CG15531*^{MI01928}*/TM3, Sb*^{1} *Ser*^{1}	*CG15531*	oxidoreductase activity	99E2-99E2
23359	*y*^{1} *w*^{67c23}; *Mi{ET1}Elo68alpha*^{MB01772}	*CG32072 - Elo68α*	fatty acid elongase activity	68A6-68A6
24819	*w*^{1118}; *Mi{ET1}Octbeta3R*^{MB04794}	*CG42244 - Octβ3R*	Octopamine receptor activity	87B15-87C1
24490	*Sod1*$^{X-39}$ *e*^{1}*/TM3, Sb*^{1} *Ser*^{1}	*CG11793 - Superoxide dismutase 1*	Superoxide dismutase	68A7-68A7
44999	*y*^{1} *w**; *Mi{y*$^{+mDint2}$*=MIC}Ir67b*^{MI08762}*/ TM3, Sb*^{1} *Ser*^{1}	*CG12303 - Ir67b*	ligand-gated ion channel	67E3-67E3
59181	*y*^{1} *w**; *Mi{y*$^{+mDint2}$*=MIC}Ir67c*^{MI13686}*/ TM3, Sb*^{1} *Ser*^{1}	*CG32058 - Ir67c*	ligand-gated ion channel	67E3-67E3
42321	*y*^{1} *w**; *Mi{y*$^{+mDint2}$*=MIC}CG14395*^{MI05388}	*CG14395*	ucharacterized	87C4-87C5
18896	*w*^{1118}; *PBac{w*$^{+mC}$*=WH}Octbeta2R*^{f05679}	*CG33976 - Octβ2R*	octopamine receptor activity	87C1-87C2

Stock[1]	Genotype[1]	Affected gene[2]	Gene function[2]	Cytological region[2]
23491	*w^{1118}; Mi{ET1}Hug^{MB02782}*	*CG6371 - Hug*	myostimulatory hormone activity; neuropeptide receptor binding	87B15-87B15
3 - Stocks that produced no interspecific offspring[3]				
16668	*y^{1} w^{67c23}; P{w^{+mC} y$^{+mDint2}$=EPgy2}Snx3^{EY05688}*	*CG6359 - Snx3*	phosphatidylinositol binding	87C3-87C3
18180	*w^{1118}; PBac{w^{+mC}=RB}Not10^{e03672}*	*CG18616 - Not10*	uncharacterized	87C3-87C3
18297	*w^{1118}; PBac{w^{+mC}=WH}CG11598^{f00150}*	*CG11598*	hydrolase activity; lipase activity	87C3-87C3
11299	*w^{1118}; PBac{w^{+mC}=PB}CG6225^{c03494}/ TM6B, Tb^{1}*	*CG6225*	metalloaminopeptidase activity	87C3-87C3
13113	*w^{1118}; P{w^{+mGT}=GT1}FoxK^{BG01104}*	*CG11799 - FoxK*	RNA pol II transcription factor	68A7-68A7
6347	*In(1)w^{m4h}; JIL-1^{3}/TM3, Sb^{1} Ser^{1}*	*CG6297 - JIL-1*	histone kinase activity	68A5-68A6

Stock[1]	Genotype[1]	Affected gene[2]	Gene function[2]	Cytological region[2]
15573	*y[1] w[67c23]; P{w[+mC] y[+mDint2]=EPgy2}CG43693[EY02590]/TM3, Sb[1] Ser[1]*	*CG7888; CG43693*	amino acid transmembrane transporter activity (both genes)	68A3-68A3; 68A2-68A2
19084	*w[1118]; PBac{w[+mC]=WH}CG32081[f07587]*	*CG32081*	amino acid transmembrane transporter activity	68A1-68A2
15006	*y[1] w[/l] ; P{w+mC y+mDint2=EPgy2}CG5946[EY00183]*	*CG5946*	oxidoreductase activity	68E1-68E1
63777	*w[1118]; PBac{w[+mC]=IT.GAL4}Oct-TyrR[0774-G4]*	*CG7485 - Oct-TyrR*	Tyramine 1 class receptor	79B1-79B2
26177	*verm[76BDx-6] red[1] e[4]/TM6B, Sb[1] Tb[1] ca[1]*	*CG8756 - verm*	Chitin-binding	76C1-76C3
11563	*P{ry[+t7.2]=PZ}Baldspot[02281] ry[506]/TM3, ry[RK] Sb[1] Ser[1]*	*CG3971 - Baldspot*	fatty acid elongase activity	73B4-73B5
21352	*y[1] w[67c23]; P{w[+mC] y[+mDint2]=EPgy2}JIL-1[EY04795]*	*CG6297 - JIL-1*	histone kinase activity	68A5-68A6
11623	*ry[506] P{ry[+t7.2]=PZ}Dlc90F[04091]/TM3, ry[RK] Sb[1] Ser[1]*	*CG12363 - Dlc90F*	dynein intermediate & light chain binding	91A2-91A2

Stock[1]	Genotype[1]	Affected gene[2]	Gene function[2]	Cytological region[2]
25491	*w*^{1118}; *Mi{ET1}Dnai2*^{MB06262}	*CG6053 - Dnai2*	uncharacterized	68D2-68D2
40225	*y*^{1} *w**; *Mi{y*$^{+mDint2}$*=MIC}Kif19A*^{MI05463}	*CG9913 - Kif19A*	ATP binding; microtubule motor activity	88A9-88A9
35507	*w**; *P{w*$^{+mC}$*=EP}Klp67A*^{322b24}*/TM6 B, Tb*^{1}	*CG10923 - Klp67A*	ATP binding; microtubule motor activity	67B2-67B2

[1] Disruption line stock numbers and genotypes from the Bloomington *Drosophila* Stock Center.

[2] Disrupted gene(s), putative gene function, and the cytological location of the candidate gene(s) in *D. melanogaster* are from Flybase (Gramates *et al.,* 2017).

[3] Only those stocks that produced interspecific offspring could be tested for their CHC content.

2.2.4 Disruption testing

Four genotypes are generated from disruption line x pure species crosses depending on the copy of chromosome 3 they receive: *Mel*/*Bal*; *Mel*/*Dis*; *Sim*/*Bal*; and *Sim*/*Dis* (*Mel*=WT *D. melanogaster* BJS copy of chromosome 3; *Sim*=WT *D. simulans* FC copy of chromosome 3; *Dis=D. melanogaster* chromosome 3 with a candidate gene disrupted; *Bal=D. melanogaster* balancer chromosome TM6b). Offspring with genotypes *Mel*/*Bal*, *Mel*/*Dis*, or *Sim*/*Bal* contain a complete copy of the *D. melanogaster* chromosome 3. The *D. melanogaster* CHC profile is semi-dominant over the *D. simulans* profile in hybrids between these species (Coyne, 1996), so flies with these genotypes should all display *melanogaster*-like CHC profiles. *Sim*/*Dis* flies contain a complete copy of the *D. melanogaster* chromosome 3 except for a single disrupted candidate gene. If the disrupted gene contributes to the dominant *D. melanogaster* CHC profile, the hybrid CHC profile should appear more *simulans*-like than the WT hybrid since the recessive *D. simulans* genotype will be unmasked at this locus. Candidate genes were tested in a version of complementation analysis, in which disruption of a particular *D. melanogaster* gene partially restores the *D. simulans* CHC phenotype if this gene is involved in species differences in CHC production and the semi-dominant *D. melanogaster* profile observed in hybrids. After accounting for differences due to species (*melanogaster* vs. hybrid) and genotype (*Dis* vs. *Bal*), I am looking for a significant species x genotype interaction.

2.2.5 Extraction and chromatography of CHC content

Virgin female flies were aged 5 days to ensure sexual maturity. Two hours after lights-on on day 5, flies were washed in 100 μL of hexane containing 100 ng of n-hexacosane internal standard; flies were gently vortexed for one minute to extract any CHCs present on the cuticle. Flies were then removed from the hexane using forceps, placed into an open Eppendorf tube and allowed to air-dry. Due to low variance in CHCs between samples of the same genotype (Moehring lab, unpublished data), only five flies from each of the four genotypes were analyzed, for a total of 20 flies assayed per line. Samples were analyzed on an Agilent Technologies (Wilmington, USA) 7890A gas chromatograph (GC), fitted

with an HP5 (5% phenyl methyl siloxane) column (30.0 m x 250.00μm internal diameter) and a flame ionization detector (at 310 °C). Samples (1 μL) were pulse-injected in splitless mode (at 200 °C with a pulse of 206 kPa for the first 1.4 min) and eluted with the following temperature program: 60°C for 0.5 min, increasing to 190°C at 120°C/min then increasing to 260°C at 7°C/min, then finally to 310°C at a rate of 120°C/min, where it was maintained for 3.5 min. Hydrogen was used as the carrier gas at a flow rate of 2.5 mL/min.

2.2.6 Analysis of CHC data

Gas chromatograph output was analyzed using Agilent Chemstation software to determine the total area of each compound that was detected. The internal standard is used as both a landmark for identifying CHCs and to correct for differences in injection volume so that different samples may be compared. Compounds were identified based on their retention times in comparison to that of the internal standard and using previous data analyzed using the same GC and parameters. Data was adjusted based on the size of the internal standard peak by multiplying the area of each peak within a sample so that the internal standard peaks across samples were equal. The average mass of flies (either pure *D. melanogaster* or hybrid *D. melanogaster/D. simulans*) was used to correct the GC output to control for possible variation in CHC abundance due to the body size of the flies. For each fly, the total peak area of each compound that was detected was scored; these values were then compared across the four genotypes.

Data were analyzed using a two-way ANOVA (α=0.05) to compare the mean areas of each compound and determine if any compounds differ significantly across the four genotypes, and to determine whether this difference is due to the interaction between genotype (*Bal* or *Dis* copy of chromosome 3), between species (*D. melanogaster* or *D. melanogaster/D. simulans* hybrid), or is the result of the genotype x species interaction. The interaction term is of greatest interest in this study as it indicates that a particular disrupted gene is implicated in species differences in CHC production. A *post hoc* assessment was performed to confirm that any significant genotype x species interactions were due to differences in the *Sim/Dis* genotype, and that changes are in the expected direction, with the amount of the compound being more similar to *D. simulans* due to

having the *D. simulans* allele unmasked. It has been suggested that the relative amounts of certain compounds may also play a role in the perception of the total CHC profile in *Drosophila* (Ferveur and Sureau, 1996; Savarit *et al.,* 1999). Therefore, in addition to comparing the total areas of each compound, relative abundance of each compound, in which the area of the peak of interest was divided by the total area of all peaks excepting that of the internal standard, was also compared. A false discovery rate (FDR) correction was used to account for multiple tests on the same data (1 ANOVA per compound, 10-12 compounds per line, 10 lines).

2.2.7 Sequence alignment

DNA sequences from *D. melanogaster* and *D. simulans* were gathered using sequences available on Flybase (Gramates *et al.,* 2017). Sequences were compared using BLASTn (for nucleotide alignment) or translated protein sequences were aligned using BLASTp (BLAST: Altschul *et al.,* 1990).

2.3 Results

2.3.1 Candidate genes that are inviable in hybrids

Some candidate genes could not be tested, either due to a total lack of hybrid production because of a lack of interspecific mating, or due to hybrids that were unable to survive to the testing age of five days. Candidate genes that, when disrupted, resulted in no larvae being produced are listed in Table 2.1, section 3. It is possible that the absence of larvae is because *D. melanogaster* females that are hemizygous for these candidate genes have a reduced propensity to mate with *D. simulans* males, and thus no mating occurred. Alternatively, mating could have occurred, but F1 offspring died as eggs and never hatched into larvae.

Some candidate gene disruption lines were successful in mating and producing offspring with *D. simulans*, but produced offspring that did not eclose from the pupal case to produce viable adults, or produced adults that consistently died within one or two days and thus did not survive to testing age (listed in Table 2.1, section 2). The candidate genes being tested in these cases must have a dramatic effect on the health of hybrid flies when no *D. melanogaster* alleles are present and only the *D. simulans* allele is able to be expressed.

2.3.2 Candidate genes in viable hybrids

Females of ten candidate gene disruption lines were able to reproduce with *D. simulans* males and produced enough offspring that lived until day five to test. The disruption lines that were successfully tested are listed in Table 2.1, section 1. All hybrid flies (*Sim/Dis* and *Sim/Bal*) produced more total (~2.5 fold more) CHCs than pure *D. melanogaster* flies (*Mel/Dis* and *Mel/Bal*), despite hybrids having a slightly smaller body size. In all candidate gene disruption lines that were tested, relative abundance and peak area comparisons both yielded a single significant result, which was for the disruption of candidate gene *CG5946*.

2.3.3 Disruption of candidate gene *CG5946*

Of the 10 disruption lines that generated enough interspecific offspring to test, one showed significance due to the species x genotype interaction in the *Sim*/*Dis* hybrids. In this stock candidate gene *CG5946* was disrupted; *Sim*/*Dis* hybrids with this disruption had significantly higher levels of 7-T and a significant decrease in 7,11-HD when quantified using either total peak area or relative abundance (Figures 2.2 and 2.3; see appendix A). Proteins that may interact with the product of *CG5946* were identified using STRINGdb (Figure 2.4). None of the remaining 9 lines showed significant differences in any of the compounds identified in either total peak area or relative abundance (Figures 2.5, 2.6 and 2.7; see appendix A).

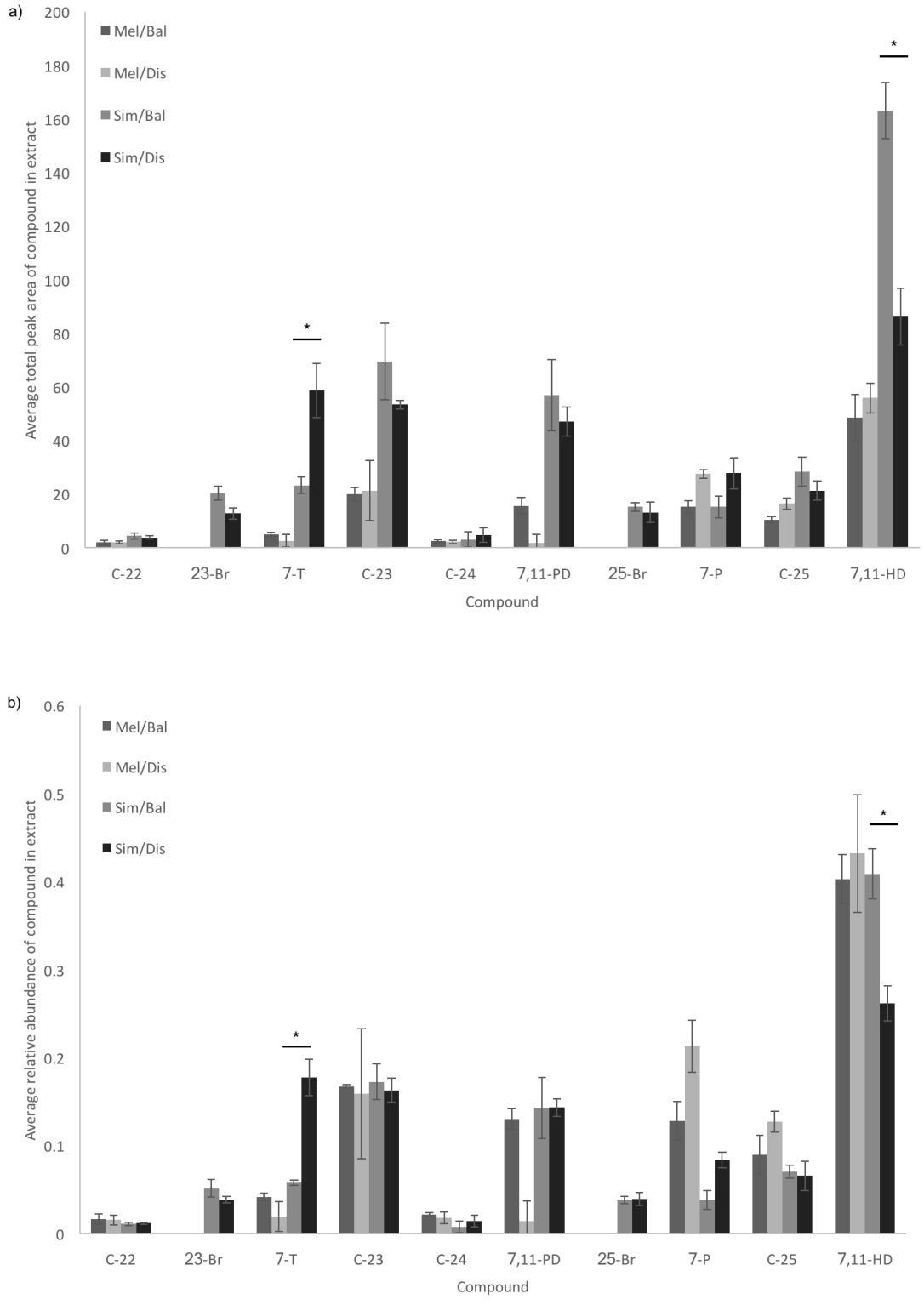


Figure 2.2: Average peak area and relative abundance of CHCs when *CG5946* is

disrupted. Bars represent the mean (±SD) total peak area or relative abundance of compounds extracted from the cuticles of 5-day old females of each genotype (*Mel/Bal*, *Mel/Dis*, *Sim/Bal*, *Sim/Dis*). a) Average peak area of each of the compounds detected across the four genotypes; b) average relative abundance of each compound across the four genotypes.

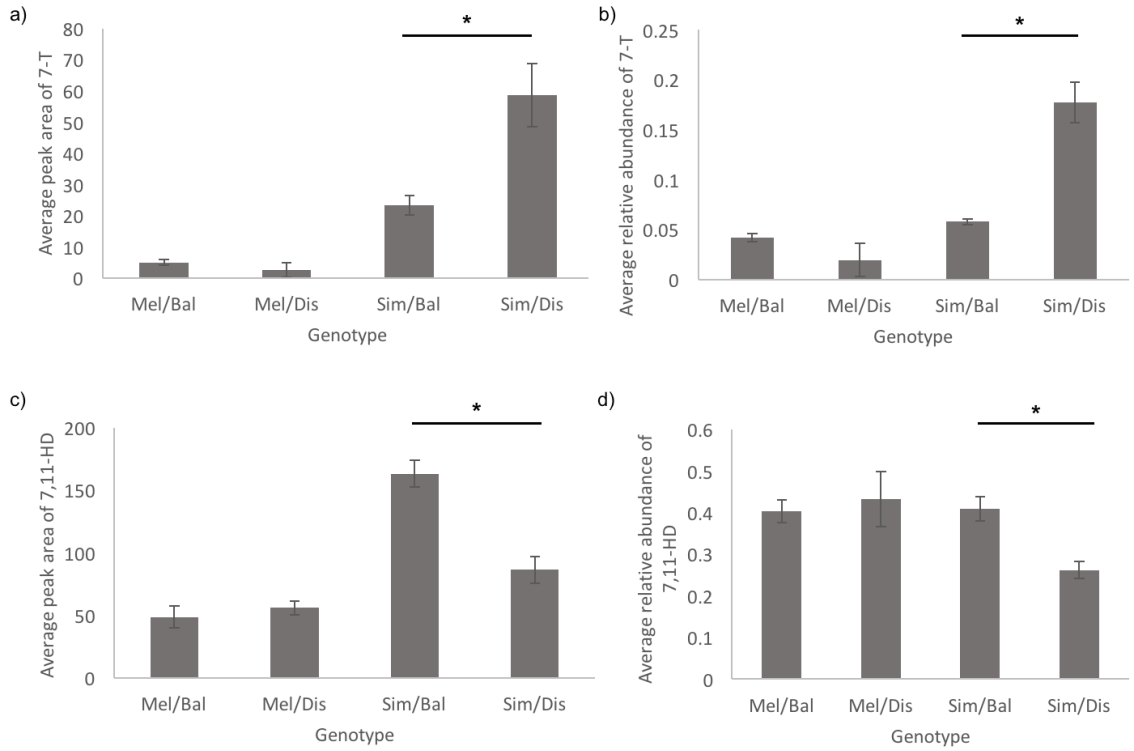
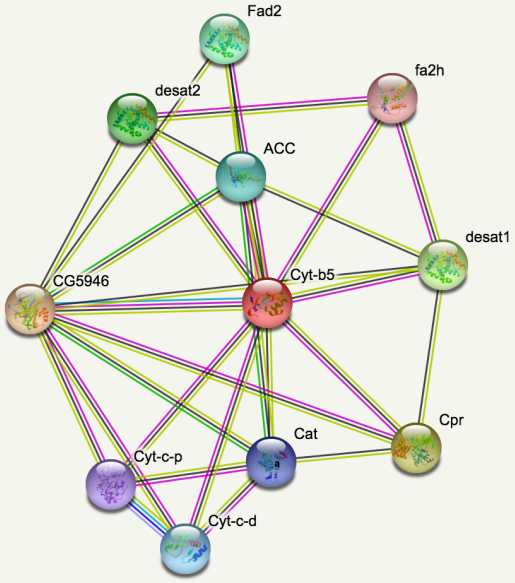


Figure 2.3: Average peak area and relative abundance of 7-T and 7,11-HD when candidate gene *CG5946* is disrupted. Bars represent the mean (±SD) total peak area or relative abundance of 7-T and 7,11-HD extracted from the cuticles of 5-day old females of each genotype (*Mel/Bal*, *Mel/Dis*, *Sim/Bal*, *Sim/Dis*). a) Average peak area of 7-T ($p_{genotype}$=7.661x10^{-4}, p_{hybrid}=2.403x10^{-6}, $p_{interaction}$=3.196x10^{-4}, two-way ANOVA, =6.046, 4 d.f.); b) Average relative abundance of 7-T ($p_{genotype}$=2.461x10^{-4}, p_{hybrid}=3.592x10^{-6}, $p_{interaction}$=1.650x10^{-5}, two-way ANOVA, t=10.08, 4 d.f.); c) average peak area of 7,11-HD ($p_{genotype}$=1.648x10^{-4}, p_{hybrid}=7.235x10^{-7}, $p_{interaction}$=4.258x10^{-5}, two-way ANOVA, t=-12.60, 4 d.f.); d) average relative abundance of 7,11-HD ($p_{genotype}$=3.490^{-2}, p_{hybrid}=7.578x10^{-3}, $p_{interaction}$=5.239x10^{-3}, two-way ANOVA, t=-12.75, 4 d.f.).

a)



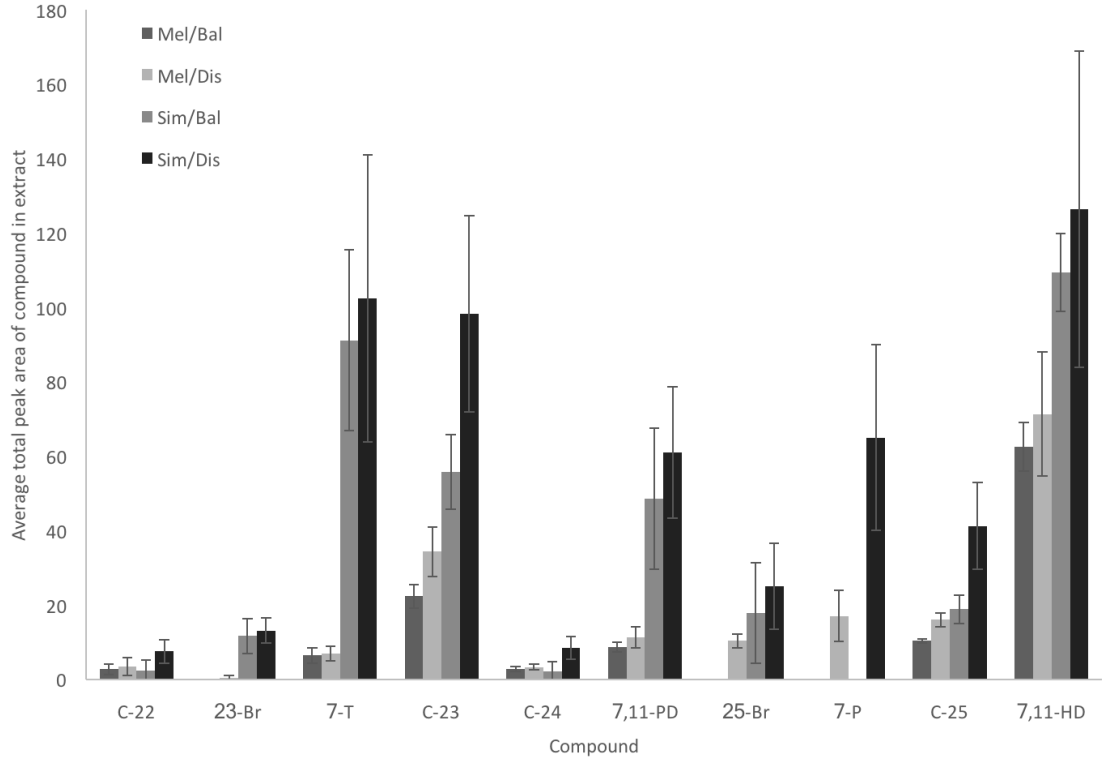
b)

Predicted Functional Partners:

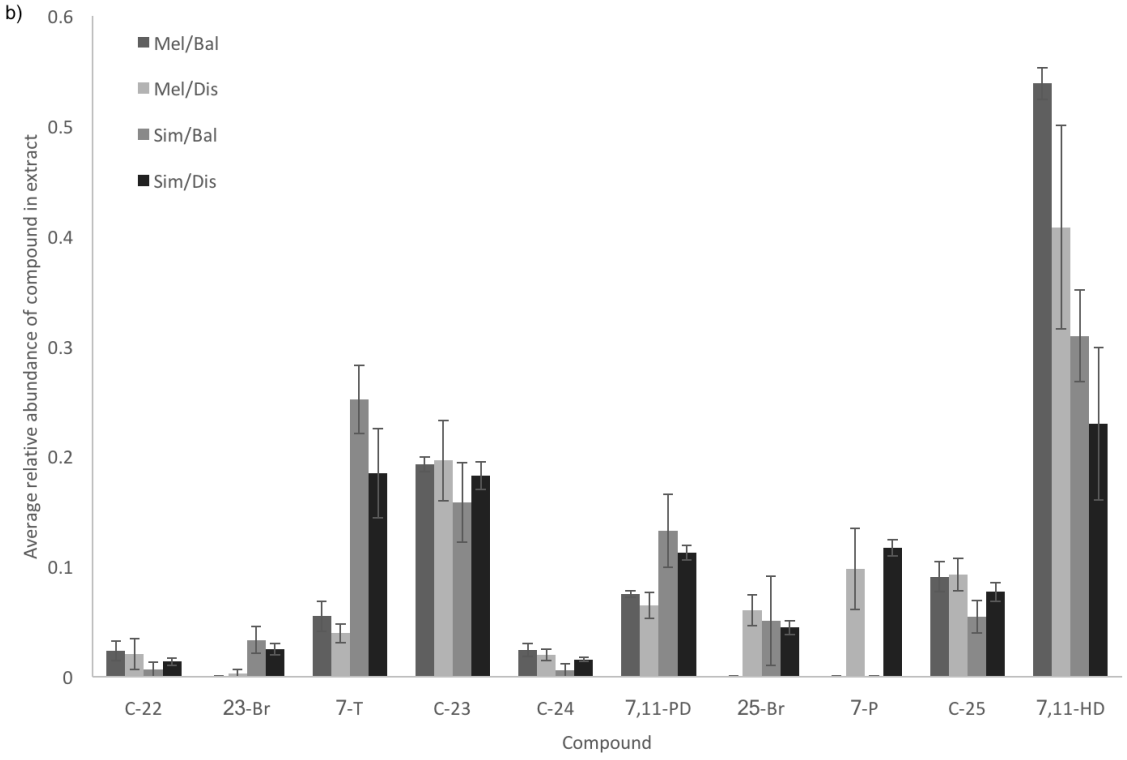
		Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
CG5946	*Cytochrome-b5 reductase activity, acting on NAD(P)H (316 aa)*						●	●		0.968
Cpr	*Cytochrome P450 reductase; This enzyme is required for electron transfer from NADP to cytochrome p450 in microsomes. It c...*							●		0.820
desat1	*Desat1 (383 aa)*							●		0.734
desat2	*Desat2 (361 aa)*							●		0.733
Fad2	*Fad2 (355 aa)*					●		●		0.732
ACC	*Acetyl-CoA carboxylase (2571 aa)*							●		0.623
Cyt-c-d	*Cytochrome c distal; Electron carrier protein. The oxidized form of the cytochrome c heme group can accept an electron from t...*							●		0.608
Cat	*Catalase; Occurs in almost all aerobically respiring organisms and serves to protect cells from the toxic effects of hydrogen pe...*							●		0.604
Cyt-c-p	*Cytochrome c proximal; Electron carrier protein. The oxidized form of the cytochrome c heme group can accept an electron fro...*							●		0.599
fa2h	*Fatty acid 2-hydroxylase; Required for alpha-hydroxylation of free fatty acids and the formation of alpha-hydroxylated sphingolip...*							●		0.557

Figure 2.4: Proteins potentially interacting with cyt-b5 and CG5946 in *D. melanogaster*. From the STRING database (Jensen *et al.,* 2009). a) Protein interaction map centered on Cyt-b5. Different coloured lines represent the different lines of evidence used to form associations between putatively interacting proteins. b) Evidence and total score for interaction between Cyt-b5 and its top 10 interacting partners.

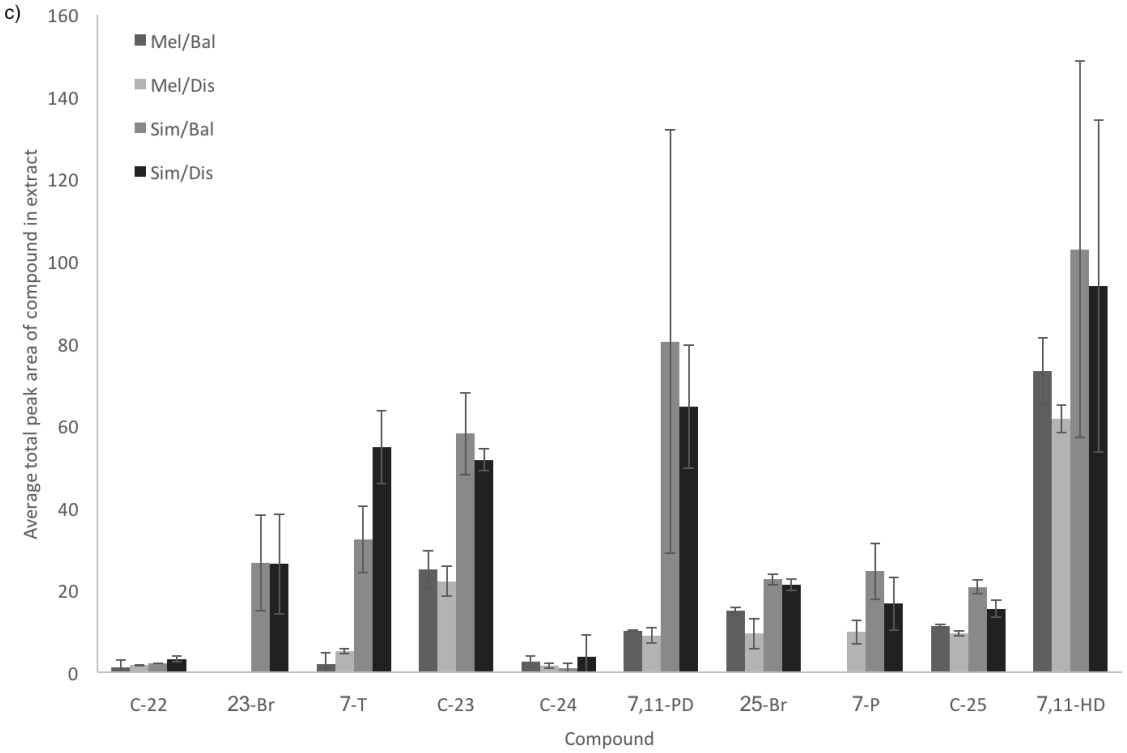
a)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average total peak area of compound in extract
0
20
40
60
80
100
120
140
160
180
C-22
23-Br
7-T
C-23
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
Compound

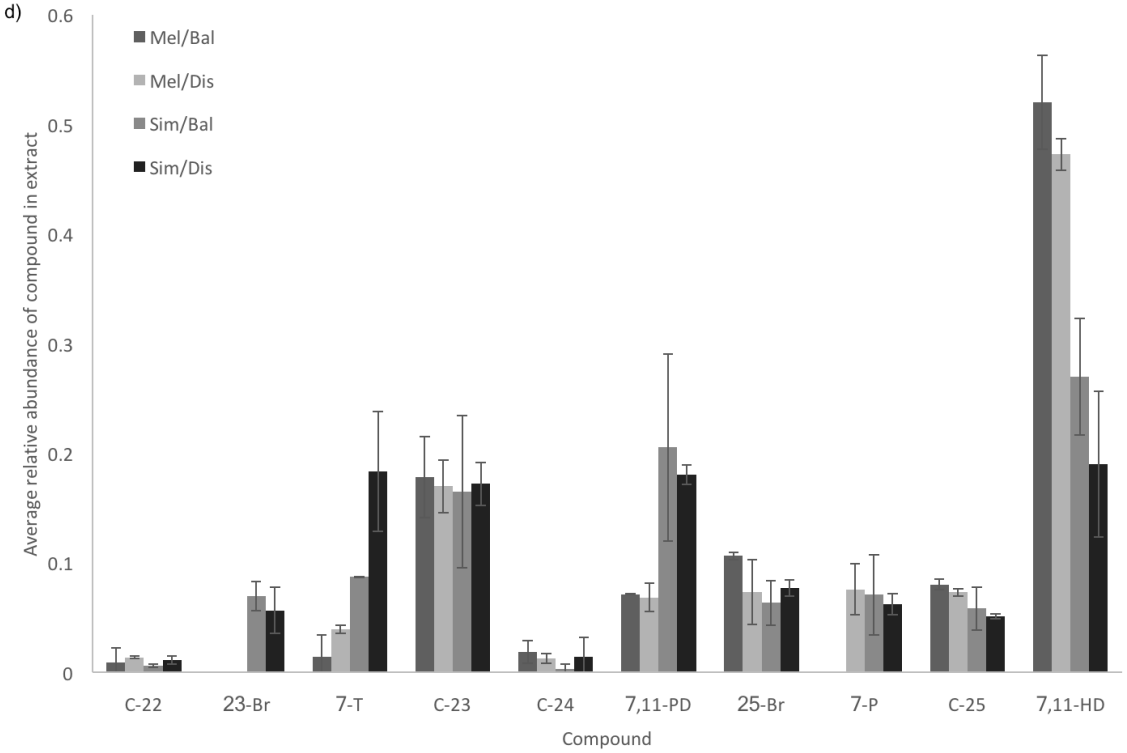
b)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average relative abundance of compound in extract
0
0.1
0.2
0.3
0.4
0.5
0.6
C-22
23-Br
7-T
C-23
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
Compound

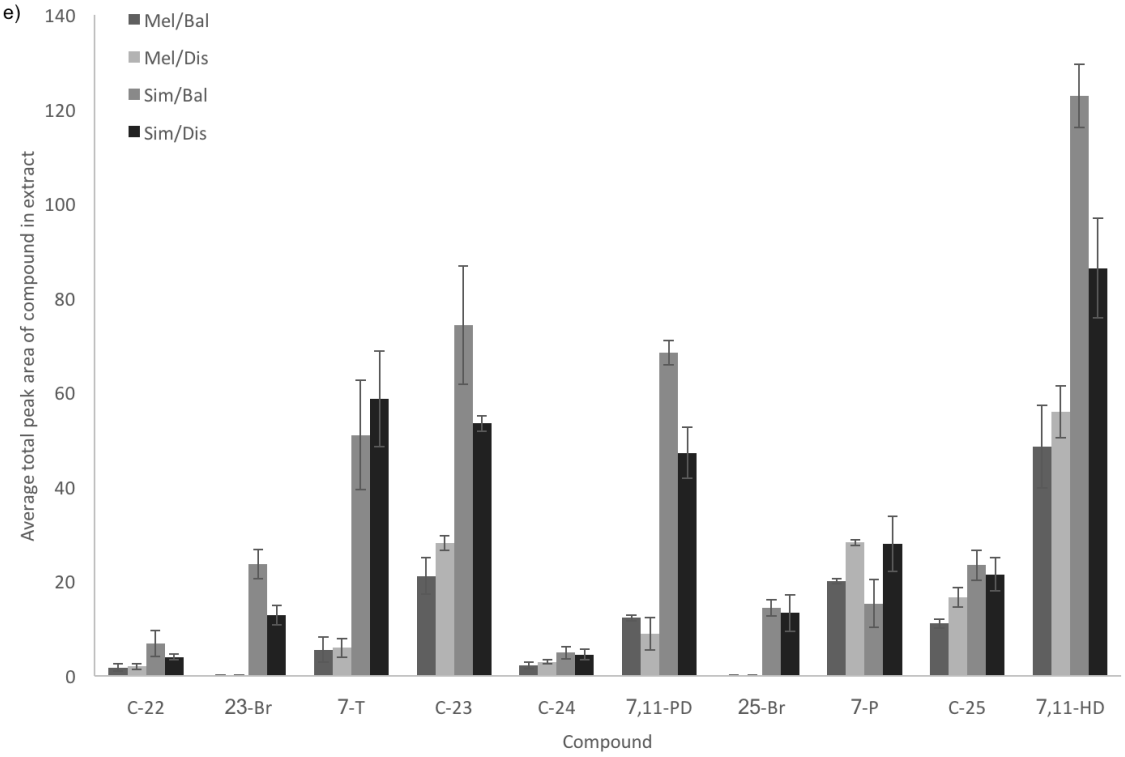
c)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average total peak area of compound in extract
0
20
40
60
80
100
120
140
160
C-22
23-Br
7-T
C-23
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
Compound

d)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average relative abundance of compound in extract
0
0.1
0.2
0.3
0.4
0.5
0.6
C-22
23-Br
7-T
C-23
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
Compound

e)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average total peak area of compound in extract
0
20
40
60
80
100
120
140
C-22
23-Br
7-T
C-23
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
Compound

f)

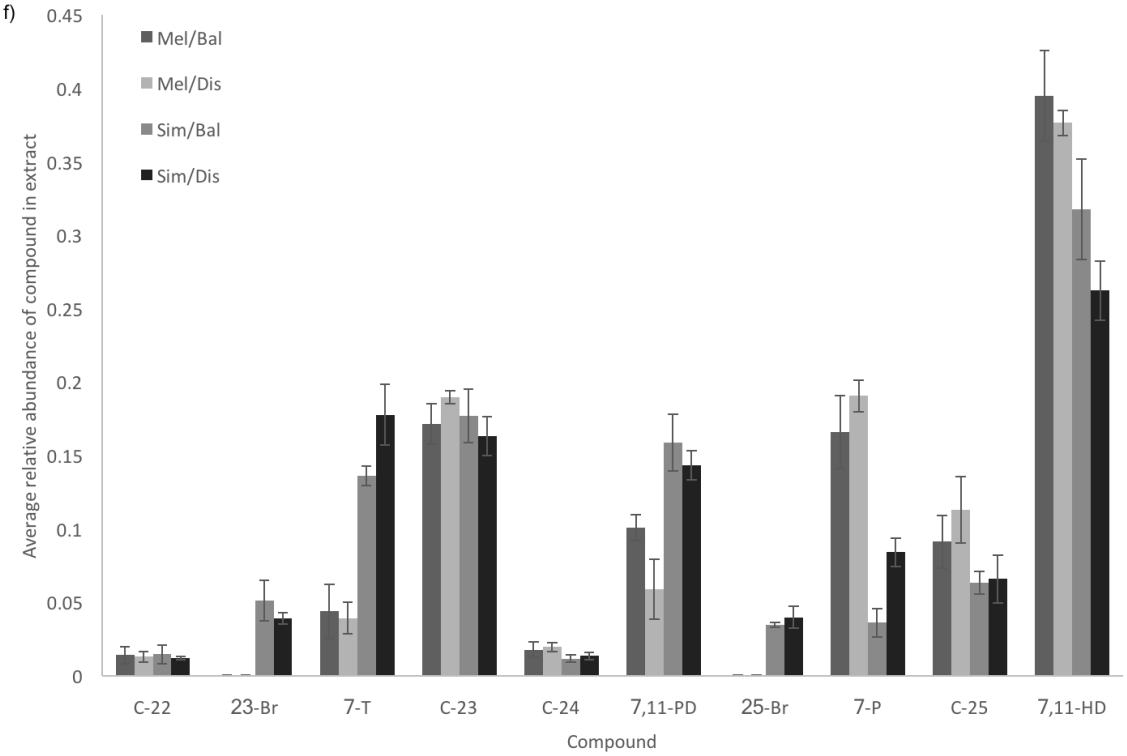
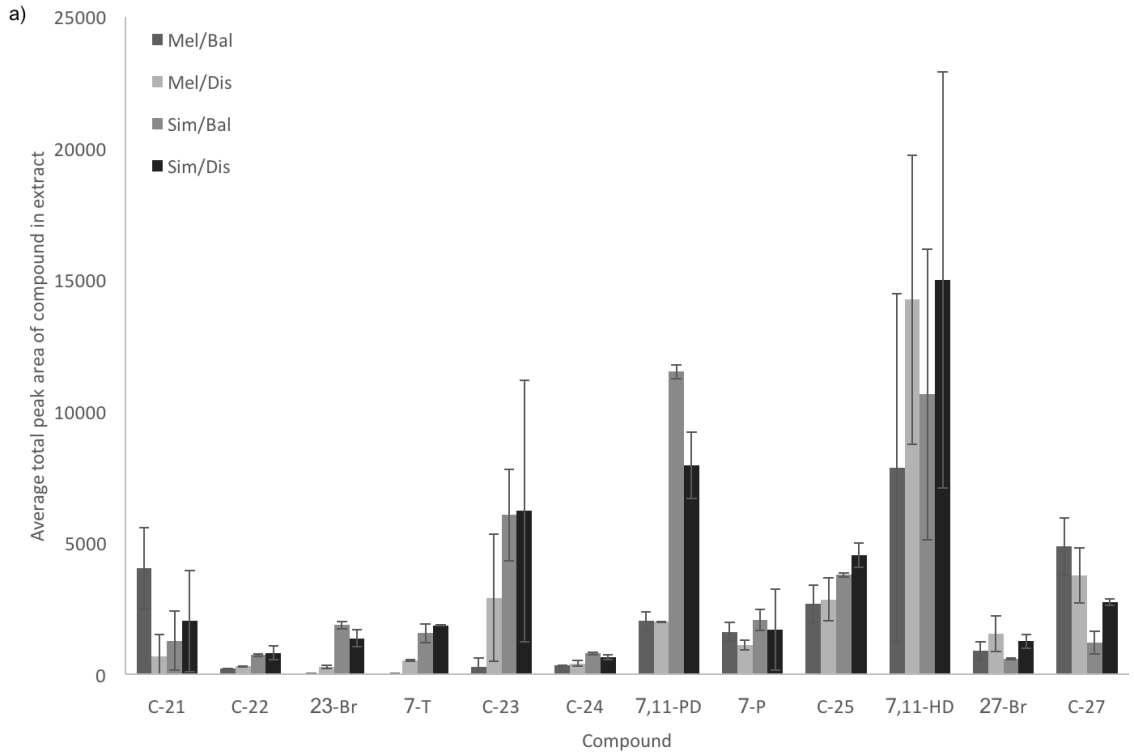
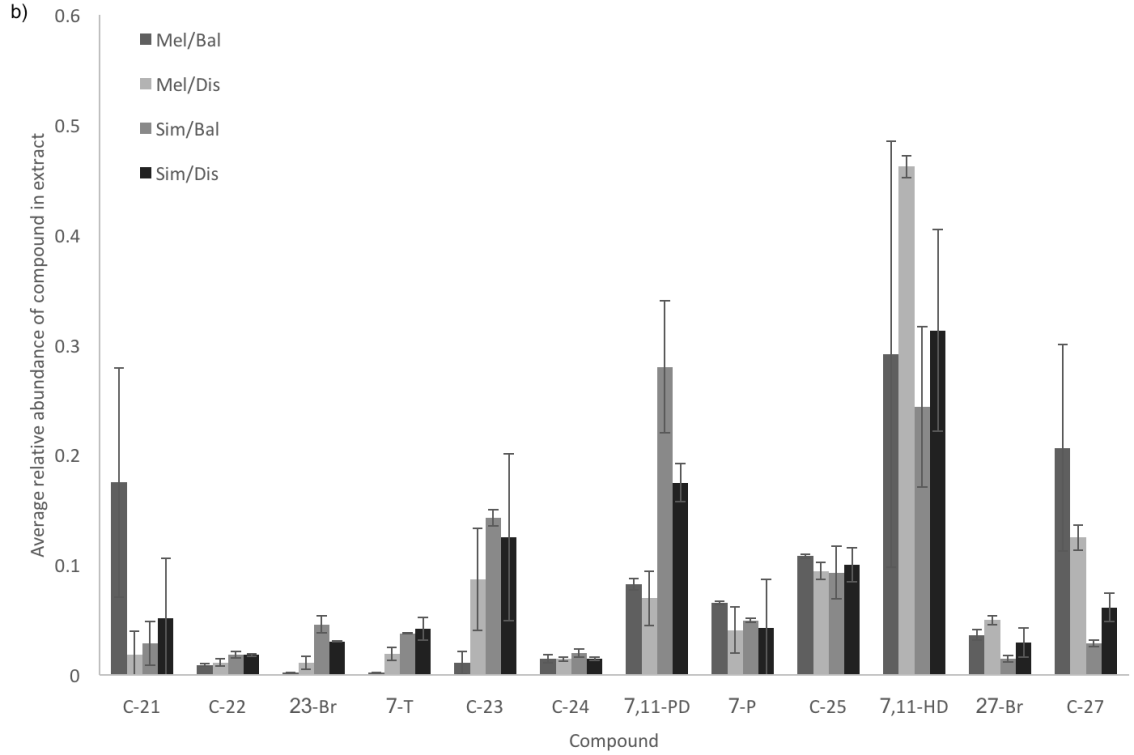
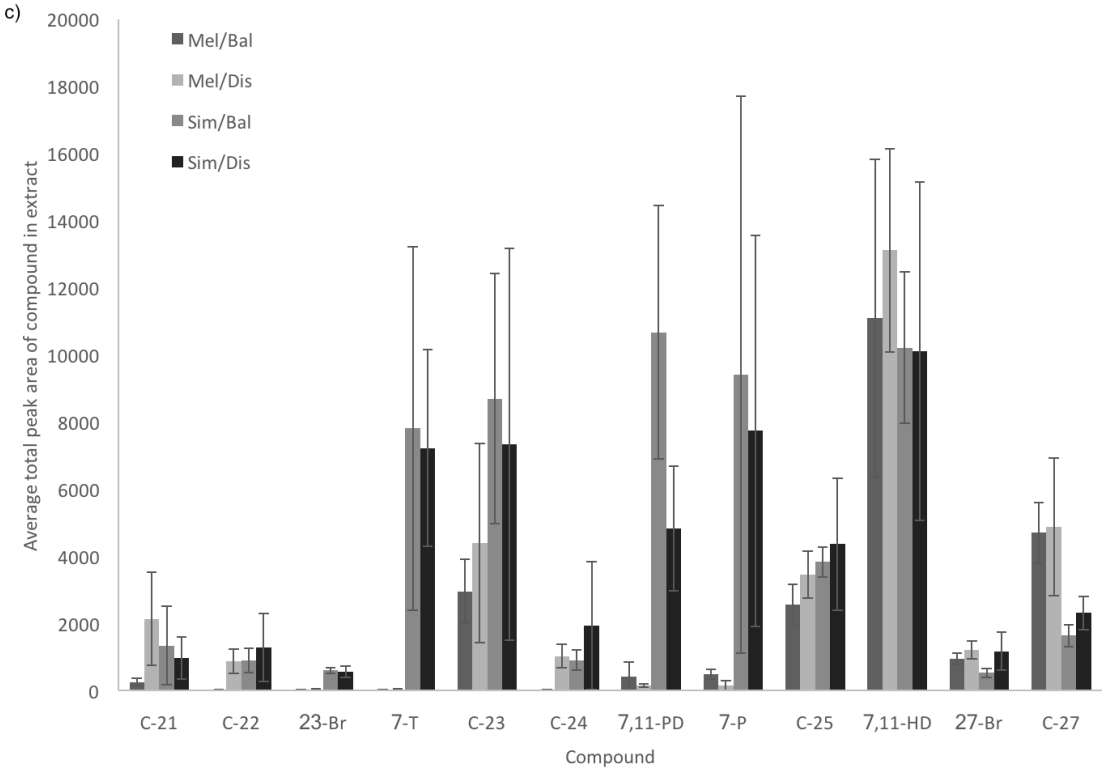

Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average relative abundance of compound in extract
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
C-22
23-Br
7-T
C-23
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
Compound

Figure 2.5: Average peak area and relative abundance of CHCs when CG1945, CG5278 or CG6921/CG44062 are disrupted. Bars represent the mean (±SD) total peak area or relative abundance of compounds extracted from the cuticles of 5-day old females of each genotype (*Mel/Bal*, *Mel/Dis*, *Sim/Bal*, *Sim/Dis*). Average total peak area (a, c, e) and average relative abundance (b, d, f) of compounds when *CG19456* (a, b) *CG5278* (c, d), or *CG6921/CG44062* (e, f) is disrupted.

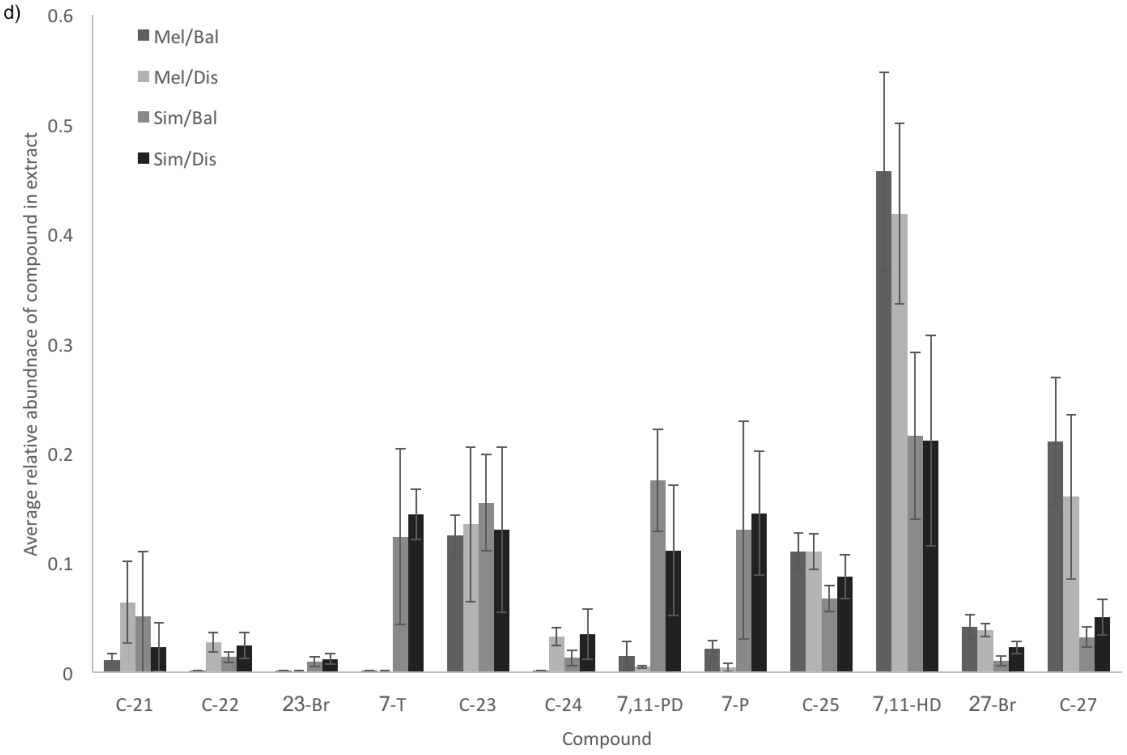

a)
25000
20000
15000
10000
5000
0
Average total peak area of compound in extract
Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
C-21
C-22
23-Br
7-T
C-23
C-24
7,11-PD
7-P
C-25
7,11-HD
27-Br
C-27
Compound


b)
0.6
0.5
0.4
0.3
0.2
0.1
0
Average relative abundance of compound in extract
Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
C-21
C-22
23-Br
7-T
C-23
C-24
7,11-PD
7-P
C-25
7,11-HD
27-Br
C-27
Compound

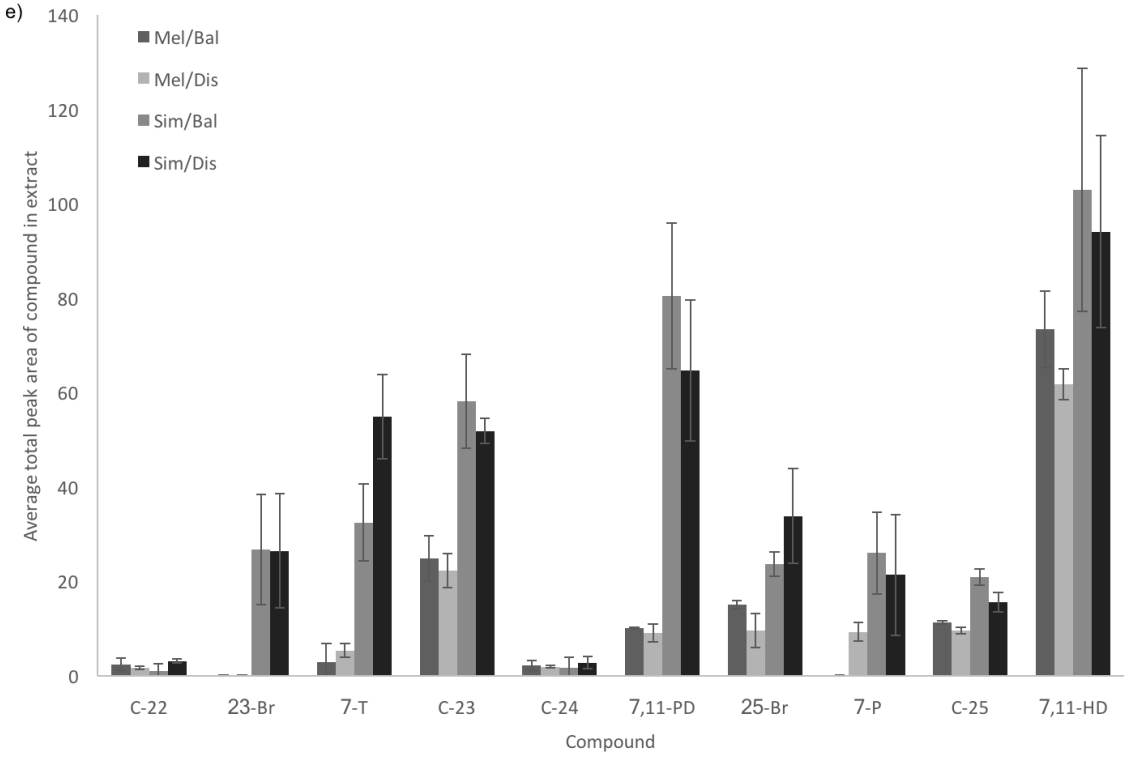
c)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average total peak area of compound in extract
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
C-21
C-22
23-Br
7-T
C-23
C-24
7,11-PD
7-P
C-25
7,11-HD
27-Br
C-27
Compound

d)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average relative abundnace of compound in extract
0.6
0.5
0.4
0.3
0.2
0.1
0
C-21
C-22
23-Br
7-T
C-23
C-24
7,11-PD
7-P
C-25
7,11-HD
27-Br
C-27
Compound

e)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average total peak area of compound in extract
0
20
40
60
80
100
120
140
C-22
23-Br
7-T
C-23
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
Compound

f)

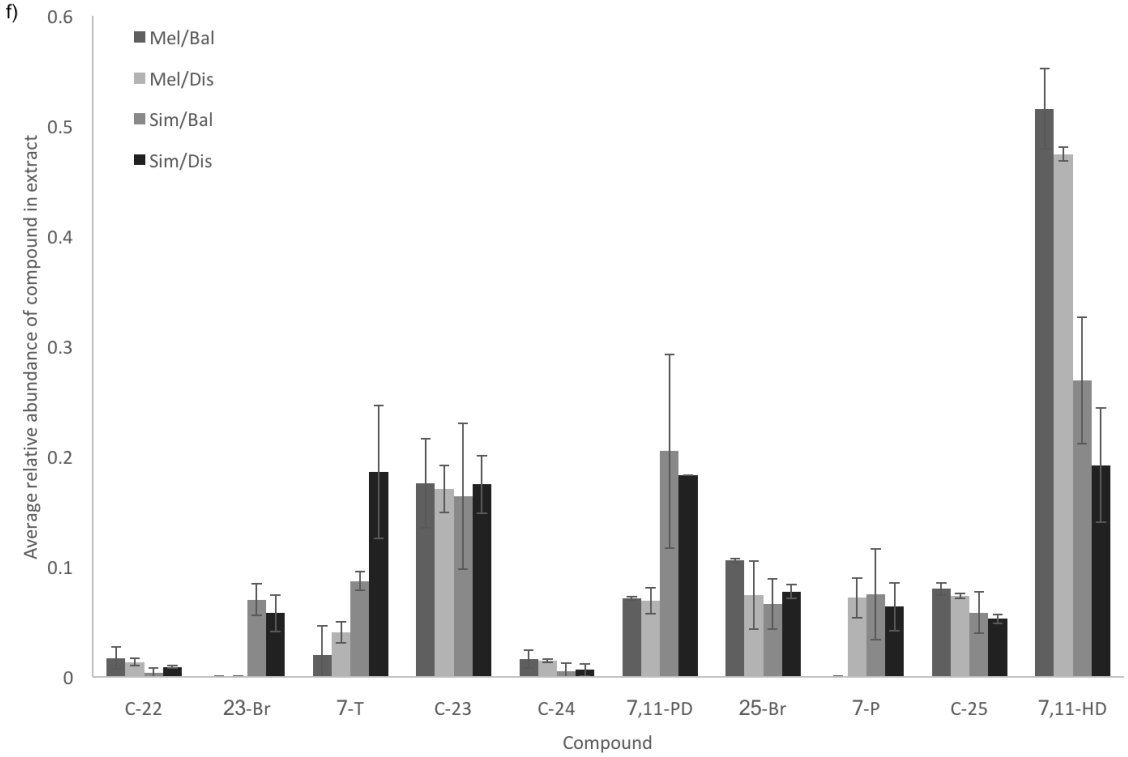
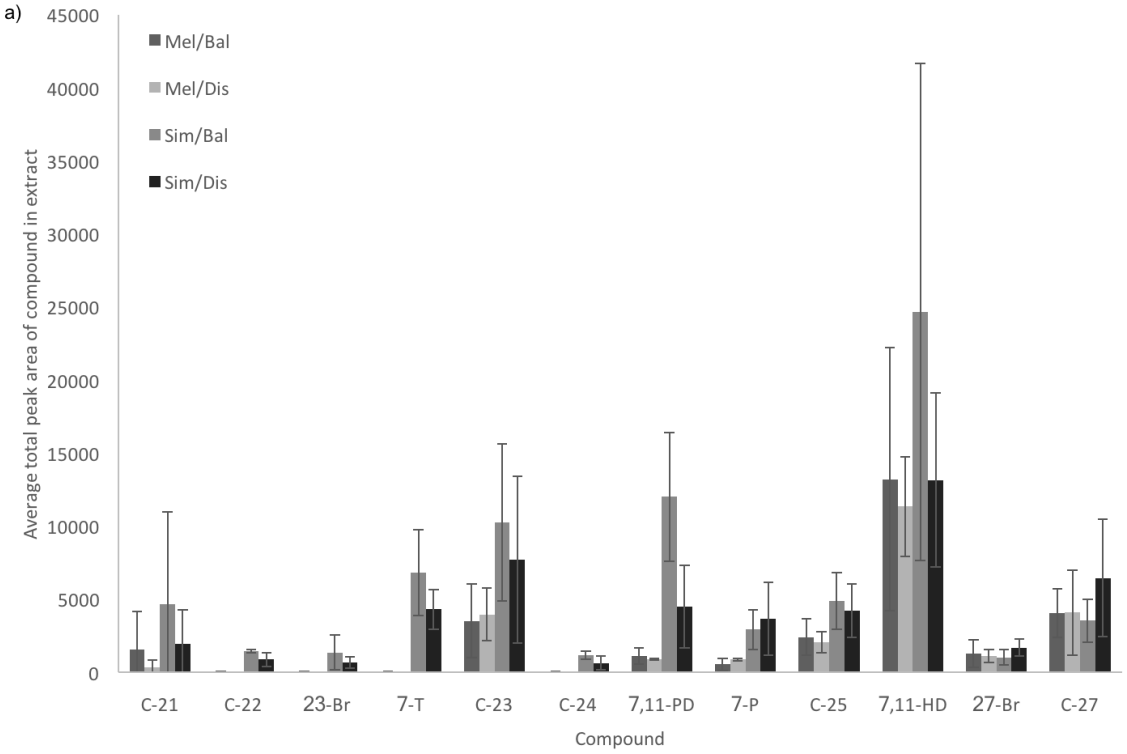
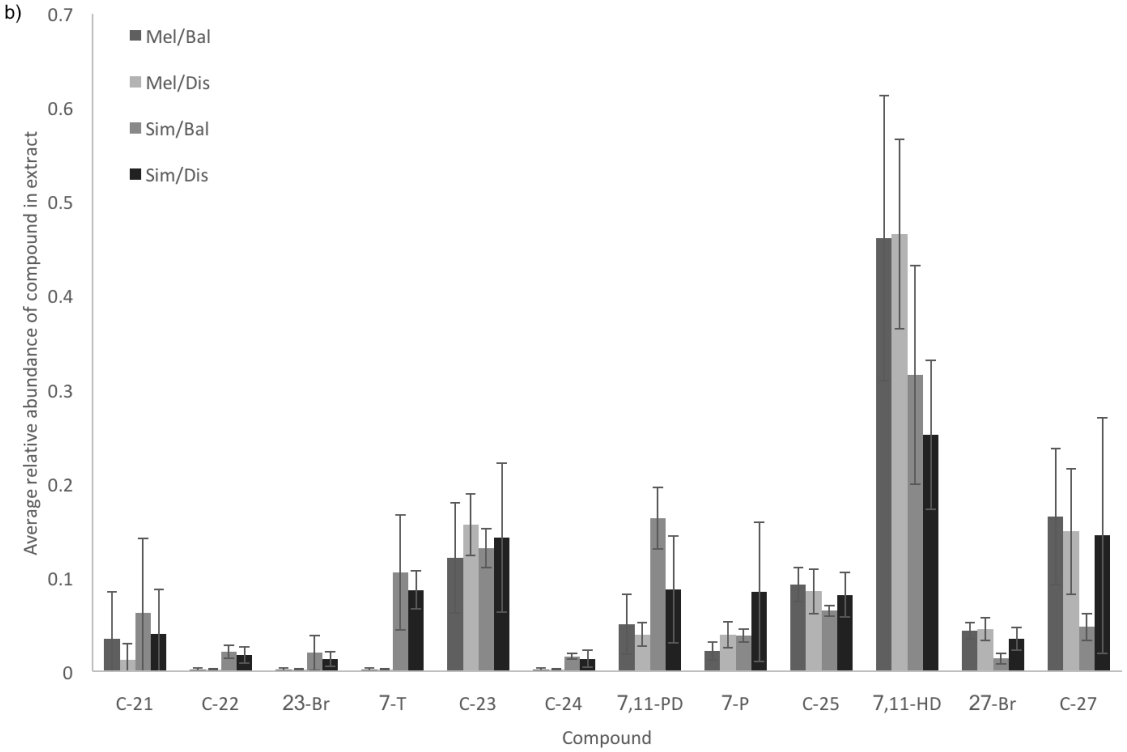

Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average relative abundance of compound in extract
0
0.1
0.2
0.3
0.4
0.5
0.6
C-22
23-Br
7-T
C-23
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
Compound

Figure 2.6: Average peak area and relative abundance of CHCs when *CG33110*, *CG42857* or *CG7485* are disrupted. Bars represent the mean (±SD) total peak area or relative abundance of compounds extracted from the cuticles of 5-day old females of each genotype (*Mel/Bal*, *Mel/Dis*, *Sim/Bal*, *Sim/Dis*). Average total peak area (a, c, e) and average relative abundance (b, d, f) of compounds when *CG33110* (a, b) *CG42857* (c, d), or *CG7485* (e, f) is disrupted.

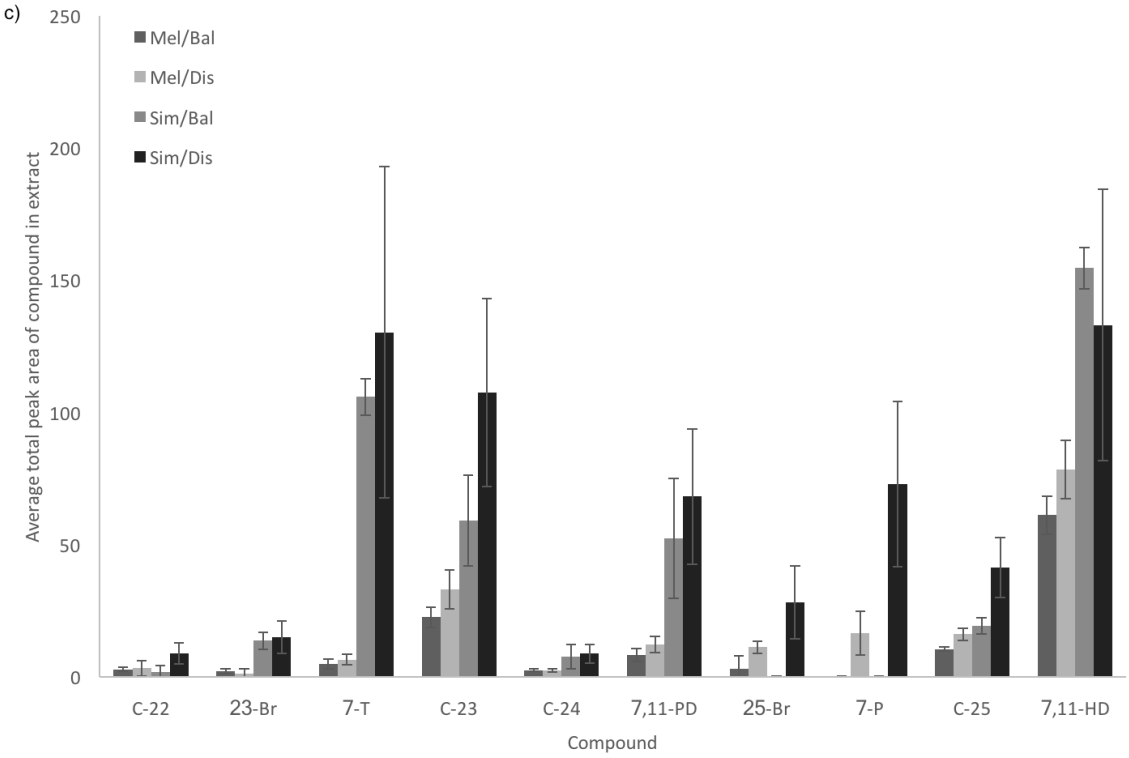
a)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average total peak area of compound in extract
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
C-21
C-22
23-Br
7-T
C-23
C-24
7,11-PD
7-P
C-25
7,11-HD
27-Br
C-27
Compound

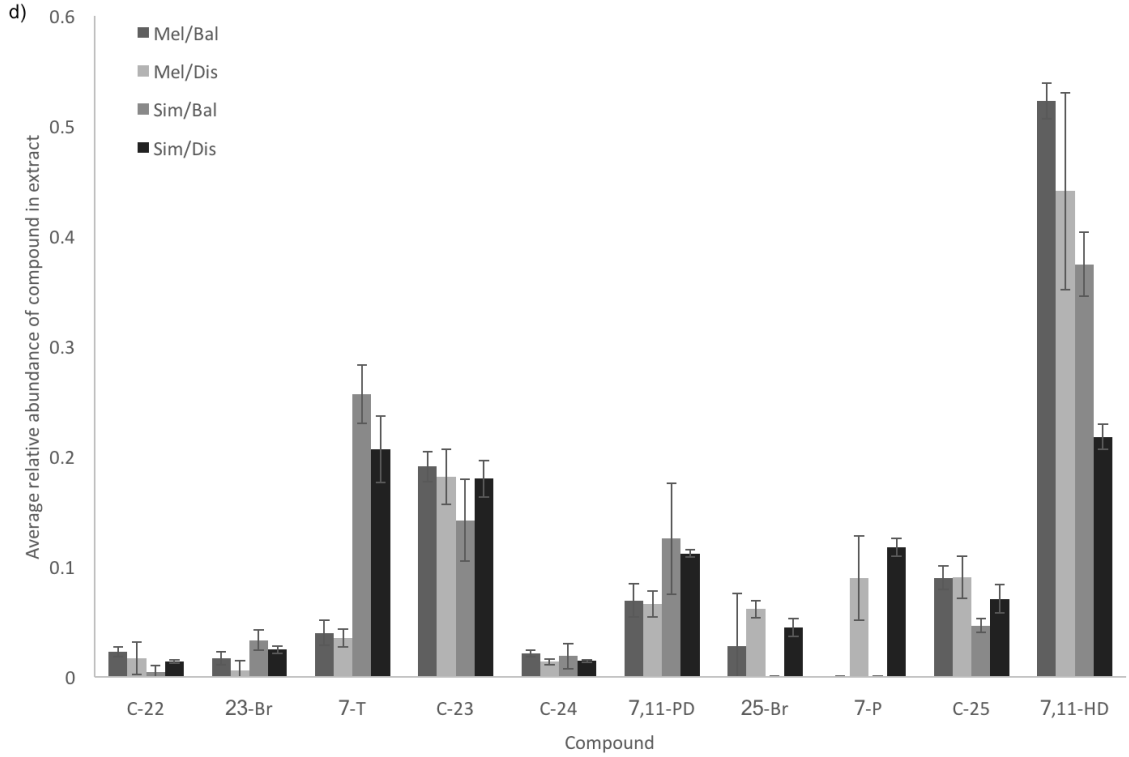
b)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average relative abundance of compound in extract
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
C-21
C-22
23-Br
7-T
C-23
C-24
7,11-PD
7-P
C-25
7,11-HD
27-Br
C-27
Compound

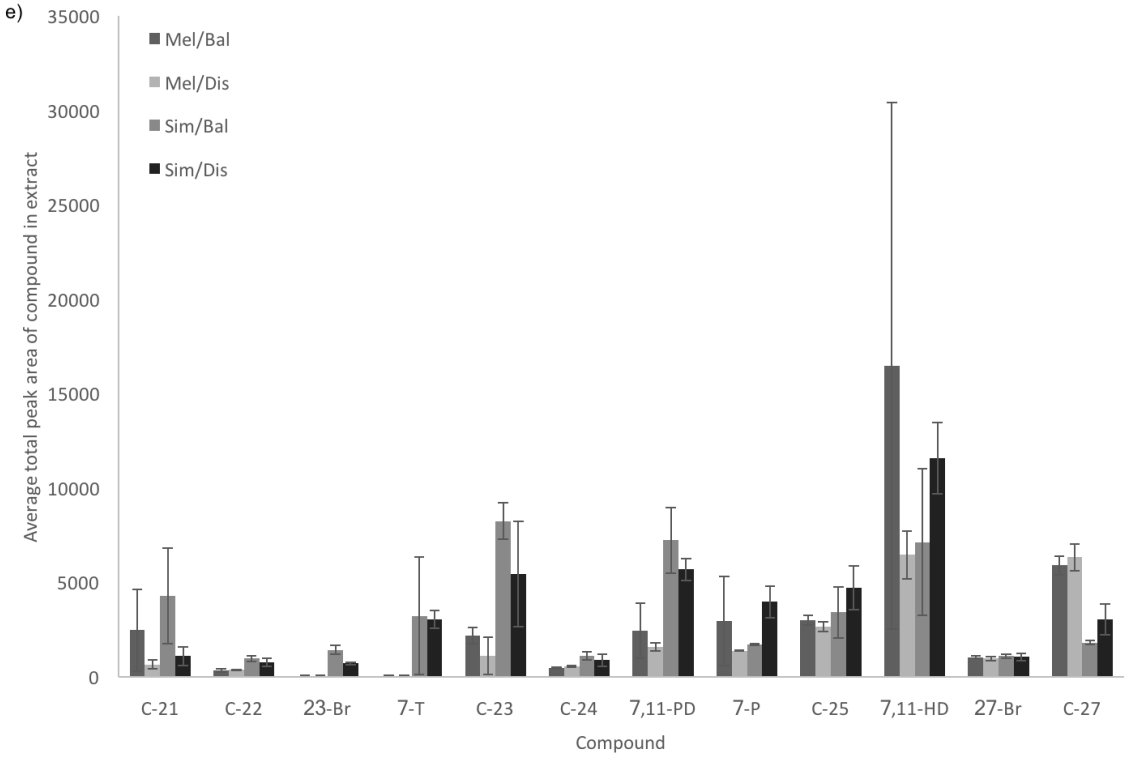
c)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average total peak area of compound in extract
250
200
150
100
50
0
C-22
23-Br
7-T
C-23
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
Compound

d)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average relative abundance of compound in extract
0.6
0.5
0.4
0.3
0.2
0.1
0
C-22
23-Br
7-T
C-23
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
Compound

e)


Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average total peak area of compound in extract
35000
30000
25000
20000
15000
10000
5000
0
C-21
C-22
23-Br
7-T
C-23
C-24
7,11-PD
7-P
C-25
7,11-HD
27-Br
C-27
Compound

f)

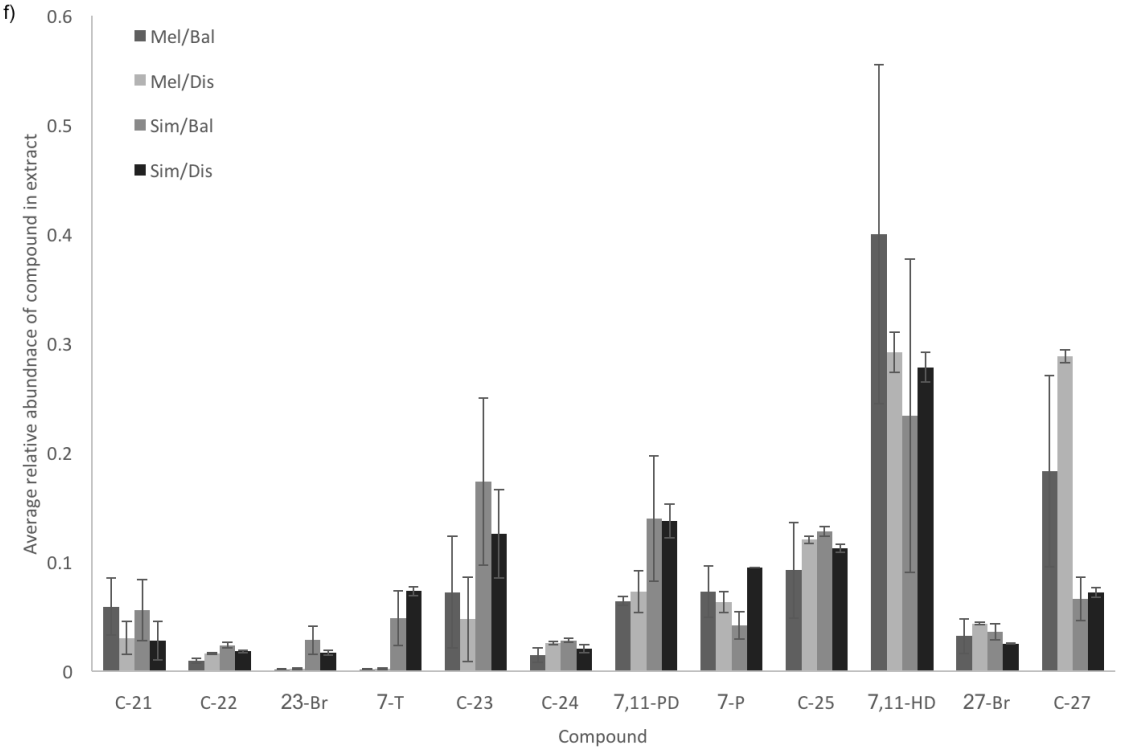

Mel/Bal
Mel/Dis
Sim/Bal
Sim/Dis
Average relative abundnace of compound in extract
0.6
0.5
0.4
0.3
0.2
0.1
0
C-21
C-22
23-Br
7-T
C-23
C-24
7,11-PD
7-P
C-25
7,11-HD
27-Br
C-27
Compound

Figure 2.7: Average peak area and relative abundance of CHCs when *CG8756*, *CG8522* or *CG3971* are disrupted. Bars represent the mean (±SD) total peak area or relative abundance of compounds extracted from the cuticles of 5-day old females of each genotype (*Mel/Bal*, *Mel/Dis*, *Sim/Bal*, *Sim/Dis*). Average total peak area (a, c, e) and average relative abundance (b, d, f) of compounds when *CG8756* (a, b) *CG8522* (c, d), or *CG3971* (e, f) is disrupted.

2.4 Discussion

In order to determine which genes contribute to CHC production and differences between *D. melanogaster* and *D. simulans*, I tested candidate genes involved in fatty acid biosynthetic processes in a version of complementation analysis. I compared the CHC profiles of hybrid offspring to determine whether any compounds shifted dramatically in abundance when hybrids lack a *D. melanogaster* copy of a particular candidate gene, resulting in the unmasking of the *D. simulans* allele at that locus. Many of the candidate gene disruption lines failed to mate or to produce enough healthy offspring to test in interspecific crosses with *D. simulans*, however 10 lines were successfully crossed to *D. simulans* and their CHCs were examined. Of the 10 candidate genes tested, the disruption of one, *CG5946*, had a significant effect on CHC production in hybrid individuals, causing a significant increase in 7-T and a significant decrease in 7,11-HD, both changes that represent a more *simulans*-like profile. This indicates that the *D. melanogaster* gene *CG5946* is partially responsible for the semi-dominance of the *D. melanogaster* over the *D. simulans* CHC profile, and its disruption allows the recessive *D. simulans* profile to show through, although disruption of this gene only partially restores the *D. simulans* CHC profile. It must be noted here that depending on the disruption line being studied, only the 10-12 most abundant compounds were able to be detected. Female *D. melanogaster* flies typically produce 53 known CHCs, though many of these are produced at very low levels (Foley *et al.,* 2007). This means that the majority of CHCs were not able to be detected; it is therefore possible that some of the lines tested may have displayed differences in some of their less abundant compounds that I was not able to assess.

The gene *CG5946* is located in the left arm of chromosome 3 at cytological position 68E1. Flies deficient in the nearby region 67E2-68A7 were found to have significantly increased levels of 7-P and decreased levels of 7,11-HD, but the region that *CG5946* falls into was not tested (Pardy, 2012).

CG5946 is predicted to code for a protein, CG5946, that has not been experimentally characterized, but whose domains have been inferred via homology. The

major functions of *CG5946* according to GO ontology are cholesterol metabolism, electron transport and fatty acid desaturation (UniProt, 2017: E1JHY0). CG5946 is a transmembrane protein that has two conserved domains: a transmembrane helical domain, and a ferredoxin reductase-type FAD-binding domain. It is predicted to be an oxidoreductase that is capable of accepting electrons from NAD(P)H and reducing cytochrome b_5 through use of an FAD cofactor (UniProt, 2017: E1JHY0). This offers a possible explanation of the observed effects on CHCs, namely, a reduction in diene levels and an increase in monoenes, that result from its deletion. In the desaturation reactions that produce monoenes and dienes, three enzymes are involved that form a complex: a desaturase, a cytochrome b_5 enzyme, and a cytochrome b_5 reductase, of which CG5946 is one (reviewed in: Shanklin and Cahoon, 1998). Electrons are passed from the cytochrome b_5 reductase to cytochrome b_5 and then to the desaturase enzyme, which oxidizes the alkane into an alkene (or the monoene into a diene) with the concurrent reduction of O_2 to $2H_2O$ using the electrons that were passed on from cytochrome b_5 (reviewed in: Borgese *et al.,* 1993). Cytochrome b_5 is able to accept electrons from a number of different cytochrome b_5 reductase enzymes, so the deletion of a single one is not enough to completely ablate the ability of the flies to produce alkenes, however it does result in a reduction of the amount of electrons being funneled to the different desaturase enzymes. Although direct interaction has yet to be shown, CG5946, via cytochrome b_5, is predicted to interact with the products of *Desat1, Desat2* and *Fad2* (also called *DesatF*; Figure 2.4), all of which are desaturases (responsible for introducing double bonds into hydrocarbons). These genes have previously been shown to be involved in *Drosophila* sex pheromone production (Dallerac *et al.,* 2000; Coyne *et al.,* 1999; Chertemps *et al.,* 2007); it is possible that CG5946 does not directly complex with any of these desaturases, but acts by passing electrons to cytochrome b_5 which then provides these desaturase enzymes with the reducing power needed to carry out their reactions, introducing double bonds into hydrocarbons. Also of note as a potential interacting partner for CG5946 (or, more likely, downstream of CG5946 via cytochrome b_5) is Acetyl-CoA carboxylase (ACC), another enzyme involved in fatty acid biosynthesis (Dembeck *et al.,* 2015).

There are seven possible transcripts of *CG5946* of various lengths that are predicted to be produced by different splice variants in *D. melanogaster*, all of which share the

transmembrane and FAD-binding domains (Gramates *et al.*, 2017; UniProt, 2017: E1JHY0). *Drosophila simulans* has a homologous gene, *GD12773*, which is predicted to produce two known and four putative transcripts (Gramates *et al*., 2017; Hu *et al.,* 2013). The gene regions of *CG5946* in *D. melanogaster* and *GD12773* in *D. simulans* share 90% sequence identity (BLAST: Altschul *et al.,* 1990). The majority of the difference between these regions is restricted to the upstream and intronic regions of the genes. The upstream region of the *D. simulans GD12773* contains one putative promoter region; *D. melanogaster CG5946* contains 2 putative promoter sequences, one of which matches exactly the *D. simulans* promoter (Neural Network Promoter Prediction: Reese, 2001).

Five of the *D. simulans* transcripts correspond to six of the *D. melanogaster* transcripts (two of the *D. melanogaster* sequences have identical coding sequences), while one *D. simulans* and one *D. melanogaster* transcript are unique to each respective species (BLAST: Altschul *et al.,* 1990). The unique *D. simulans* transcript is translated into a peptide that differs only from other *D. melanogaster* and *D. simulans* transcripts in the N-terminal region (the remainder of the peptide is identical or shares 98% sequence identity with several other *D. simulans* and *D. melanogaster* peptides, but this combination of the 8 amino acid N-terminal sequence combined with the remainder of the peptide is not found in *D. melanogaster*). Since the N-terminal region is involved in targeting polypeptides to the ER, it is possible that the resultant protein is directed to a different subcellular location following translation (Lodish *et al.,* 2000). The unique *D. melanogaster* transcript is translated into a peptide which is truncated at the C-terminal end compared to other possible transcripts, but still contains the catalytic and transmembrane domains, so is likely still functional (Gramates *et al.*, 2017).

The translated proteins of the different transcripts are 98% identical to their homologues, containing only 2-5 non-synonymous mutations (Gramates *et al.,* 2017). All *D. simulans* transcript variants share the I51V (isoleucine to valine: *D. melanogaster* to *D. simulans*, counted from the start of the FAD binding domain) and S78T (serine to threonine) substitutions within the putative FAD binding domain of the protein. I51V occurs two positions before the conserved active site R-x-Y-[ST] domain, which is required for hydrogen bonding with the FAD cofactor. Despite these changes, protein structure and

activity are not likely to change since both of these substitutions are found in FAD binding domains of other species' proteins, and do not represent a change in charge or hydrophobic interactions in the proteins, and are thus unlikely to alter structure and function (reviewed in: Dym and Eisenberg, 2001).

Since the activity of the protein is likely the same in *D. melanogaster* and *D. simulans*, the changes in CHC abundance observed in the hybrids is most likely the result of differential expression in *D. simulans* and *D. melanogaster*. This is supported by the presence of an additional predicted upstream promoter sequence in *D. melanogaster* that is not present upstream of the *D. simulans* gene (Neural Network Promoter Prediction: Reese, 2001). This differential expression could be due to differences in where the genes are expressed, the exact timing of expression, or in the expression levels within tissues. Expression data does not exist for *D. simulans*, but in *D. melanogaster CG5946* has been found to be highly expressed in cells of the proboscis, ejaculatory duct, fat body, heart, spermathecae, head, oviduct, and the second antennal segment (Genevestigator: Hruz *et al.,* 2008; Bgee: Bastian *et al.,* 2008), although detailed information about individual transcript splice variant expression is lacking. It is possible that different transcripts are targeted to different functions depending where they are expressed. Although transcripts expressed in the fat body are likely solely used for fatty acid biosynthesis based on the role of cytochrome b_5 reductases in fatty acid biosynthetic processes generally (reviewed in: Borgese *et al.,* 1993), another cytochrome b_5 reductase *Cpr* which, alongside *CG5946*, is expressed in the antennae, has been suggested to have a role in clearing the antennae and preventing the accumulation of chemical signals, allowing the fly to receive new signals again more quickly (Hovemann *et al.,* 1997). Since both *Cpr* and *CG5946* act to pass electrons onto cytochrome b_5, it is possible that CG5946 is used in a similar role in the antennae. This is especially interesting considering the role of the antennae in sensing pheromones (via olfactory receptors located on the antennae: Zhuang *et al.,* 2016); if a single gene could act both to alter the biosynthesis and perception of chemical signals, this could represent a good candidate for a gene implicated in the formation of prezygotic isolating barriers. This dual-role for genes involved in both the expression and perception of a trait has been noted before, as in *Desat1* where changes in the expression of this gene can affect both the perception and expression of certain CHCs (Bousquet *et al.,* 2012).

Since CG5946 may act upstream of Desat1, this could represent an intriguing line of inquiry into whether or not CG5946 may also have effects both on perception and expression of CHCs. The antennal expression of *CG5946* could also explain the findings of Edwards *et al.* in 2009, who discovered that disruption of *CG5946* resulted in significantly decreased inter-male aggression scores; it is possible that this decreased aggression is due in part to a reduced ability to detect other males via smell once the antennae become 'clogged' and clearing is reduced due to the lack of CG5946.

In conclusion, I showed that the *D. melanogaster* gene *CG5946* is implicated in species differences in the production of the most abundant sex pheromones in hybrids between *D. melanogaster* and *D. simulans*. *CG5946* therefore represents a gene that may be responsible for some of the prezygotic reproductive barriers that exist between these species, since it contributes to the *D. melanogaster* pheromone profile that is partially responsible for the behavioural isolation observed between these species.

2.5 Bibliography

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. **1990**. Basic local alignment search tool. *Journal of Molecular Biology*. **215**:403-410.

Amrein, H. and Thorne, N. **2005**. Gustatory perception and behaviour in *Drosophila melanogaster*. *Current Biology*. **15**:673-684.

Antony, C. and Jallon, J-M. **1982**. The chemical basis for sex recognition in *Drosophila melanogaster*. *Journal of Insect Physiology*. **28**:873-880.

Bastian, F., Parmentier, G., Roux, J., Moretti, S., Laudet, V. and Robinson-Rechavi, M. **2008**. Bgee: Integrating and comparing heterogeneous transcriptome data among species. In *Data Integration in Life Sciences*. **5109**:124-131.

Billeter J.C., Atallah J., Krupp J.J., Millar J.G. and Levine J.D. **2009**. Specialized cells tag sexual and species identity in *Drosophila melanogaster*. *Nature*. **461**:987–991.

Blomquist, G.J. and Bagnères, A-G. **2010**. Insect hydrocarbons: Biology, biochemistry, and chemical ecology. Cambridge: Cambridge University Press.

Bontonou, G. and Wicker-Thomas, C. 2014. Sexual communication in the *Drosophila* genus. *Insects*. **5**:439-458.

Borgese, N., D'Arrigo, A., De Silvestris, M. and Pietrini, G. **1993**. NADH-cytochrome b5 reductase and cytochrome b5 isoforms as models for the study of post-translational targeting to the endoplasmic reticulum. *Federation of European Biochemistry Societies*. **325**:70-75.

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

Chertemps, T., Duportets, L., Labeur, C., Ueda, R., Takahashi, K., Saigo, K. et al. **2007**. A female-biased expressed elongase involved in long-chain hydrocarbon biosynthesis and courtship behavior in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA*. **104**:4273–4278.

Cobb, M., and Jallon, J-M. **1990**. Pheromones, mate recognition and courtship stimulation in the *Drosophila melanogaster* species sub-group. *Animal Behaviour*. **39**:1058-1067.

Coyne, J., Mah, K. and Crittenden, A. **1994**. Genetics of a pheromonal difference contributing to reproductive isolation in *Drosophila*. *Science*. **265**:1461-1464.

Coyne, J.A. **1996**. Genetics of differences in pheromonal hydrocarbons between *Drosophila melanogaster* and *D. simulans*. *Genetics*. **143**:353-364.

Coyne, J.A., Wicker-Thomas, C. and Jallon, J-M. **1999**. A gene responsible for a cuticular hydrocarbon polymorphism in *Drosophila melanogaster*. *Genetic Research*. **73**: 189–203.

Dallerac, R., Labeur, C., Jallon, J-M., Knippie, D.C., Roelofs, W.L. and Wicker-Thomas, C. **2000**. A Δ9 desaturase gene with a different substrate specificity is responsible for the cuticular diene hydrocarbon polymorphism in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA*. **97**:9449–9454.

Dembeck, L.M., Böröczky, K., Huang, W., Schal, C., Anholt, R.R. and Mackay, T.F. **2015**. Genetic architecture of natural variation in cuticular hydrocarbon composition in *Drosophila melanogaster*. *eLife*. **4**:e09861.

Dym, O. and Eisenberg, D. **2001**. Sequence-structure analysis of FAD-containing proteins. *Protein Science*. **10**:1712-1728.

Edwards, A.C., Zwarts, L., Yamamoto, A., Callaerts, P., Mackay, T.F. **2009**. Mutations in many genes affect aggressive behaviour in *Drosophila melanogaster*. *BMC Biology*. **7**:29.

Ellis, L.L. and Carney, G.E. **2009**. *Drosophila melanogaster* males respond differently at the behavioural and genome-wide levels to *Drosophila melanogaster* and *Drosophila simulans* females. *Journal of Evolutionary Biology*. **22**:2183-2191.

Fernández, M.P. and Kravitz, E.A. **2013**. Aggression and courtship in *Drosophila*: pheromonal communication and sex recognition. *Journal of Comparative Physiology A*. **199**:1065-1076.

Ferveur, J.F. and Sureau, G. **1996**. Simultaneous influence on male courtship of stimulatory and inhibitory pheromones produced by live sex-mosaic *Drosophila melanogaster*. *Proceedings of the Biological Society*. **22**:967-973.

Foley, B., Chenoweth, S.F., Nuzhdin, S.V. and Blows, M.W. **2007**. Natural genetic variation in cuticular hydrocarbon expression in male and female *Drosophila melanogaster*. *Genetics*. **175**: 1465-1477.

Gramates LS, Marygold SJ, dos Santos G, Urbano J-M, Antonazzo G, Matthews BB, Rey AJ, Tabone CJ, Crosby MA, Emmert DB, Falls K, Goodman JL, Hu Y, Ponting L, Schroeder AJ, Strelets VB, Thurmond J, Zhou P and the FlyBase Consortium. **2017**. FlyBase at 25: looking to the future. *Nucleic Acids Research*. **45**:D663-D671.

Hovemann, B.T., Sehlmeyer, F. and Malz, J. **1997**. *Drosophila melanogaster* NADPH-cytochrome P450 oxidoreductase: pronounced expression in antennae may be related to odorant clearance. *Genetics*. **189**:213-219.

Howard, R.W. and Blomquist, G.J. **2005**. Ecological, behavioural, and biochemical aspects of insect hydrocarbons. *Annual Review of Entomology*. **50**:371-393.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W. and Zimmermann, P. **2008**. Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes. *Advances in Bioinformatics*. 420747. https://genevisible.com/tissues/DM/UniProt/E1JHY0

Hu, T.T, Eisen, M.B., Thornton, K.R. and Andolfatto, P. **2013**. A second-generation assembly of the *Drosophila simulans* genome provides new insights into patterns of lineage-specific divergence. *Genome Research*. **23**:89-98.

Jallon, J-M. **1984**. A few chemical words exchanged by *Drosophila* during courtship and mating. *Behavioural Genetics*. **14**:441-478.

Jensen, L.J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., Bork, P. and Von Mering, C. **2009**. String 8—a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Research*. 37.

Joseph, R.M. and Carlson, J.R. **2015**. *Drosophila* chemoreceptors: A molecular interface between the chemical world and the brain. *Trends in Genetics*. **31**:683-695.

Lodish, H., Berk, A. and Zipursky, S.L. **2000**. A single internal topogenic sequence directs insertion of some single-pass transmembrane proteins. In Molecular Cell Biology, 4ed. New York: W.H. Freeman.

Matute, D.R., Gavin-Smyth, J. and Liu, G. **2014**. Variable post-zygotic isolation in *Drosophila melanogaster/D. simulans* hybrids. *Journal of Evolutionary Biology*. **27**:1691-1705.

Pardy, J. **2012**. The genetic basis of cuticular hydrocarbon production in *Drosophila melanogaster* and *D. simulans*. Electronic Thesis and Dissertation Repository. 832.

Pasyukova, E.G., Vieira, C. and Mackay, T.F. **2000**. Deficiency mapping of quantitative trait loci affecting longevity in *Drosophila melanogaster*. *Genetics*. **156**:1129-1146.

Reese, M.G. **2001**. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Computational Chemistry*. **26**:51-56.

Rouault, J.D., Marican, C., Wicker-Thomas, C. and Jallon, J-M. **2004**. Relations between cuticular hydrocarbon (HC) polymorphism, resistance to desiccation and breeding temperature; a model for HC evolution in *D. melanogaster* and *D. simulans*. *Genetica*. **120**:195-212.

Savarit F., Sureau G., Cobb M. and Ferveur J.F. **1999**. Genetic elimination of known pheromones reveals the fundamental chemical bases of mating and isolation in *Drosophila*. *Proceedings of the National Academy of Sciences USA*. **96**:9015–9020.

Shanklin, J. and Cahoon, E.B. **1998**. Desaturation and related modifications of fatty acids. *Annual Review of Plant Physiology and Plant Molecular Biology*. **49**:611-641.

Sousa-Neves, R. and Rosas, A. **2010**. An analysis of genetic changes during the divergence of *Drosophila* species. *Public Library of Science*. **5**:e10485.

Thistle, R., Cameron, P., Ghorayshi, A., Dennison, L. and Scott, K. **2012**. Contact chemoreceptors mediate male-male repulsion and male-female attraction during *Drosophila* courtship. *Cell*. **149**:1140-1151.

The UniProt Consortium. **2017**. UniProt: the universal protein knowledgebase. *Nucleic Acids Research*. **45**:D158-D169.

Yamamoto, M.T. **1992**. Inviability of hybrids between *D. melanogaster* and *D. simulans* results from the absence of *simulans* X not the presence of *simulans* Y chromosome. *Genetica*. **87**:151-158.

Zhuang, L., Sun, Y., Hu, M., Chenxi, W., La, X., Chen, X., Feng, Y., Wang, X., Hu, Y. and Xue, L. **2016**. Or47b plays a role in *Drosophila* males' preference for younger mates. *Public Library of Science*. **6**:160086

Chapter 3

3 Environmental influence on CHCs in *Drosophila melanogaster*

3.1 Introduction

Mate recognition and discrimination against heterospecifics is essential for the maintenance of species when distinct species come into contact, especially if hybrids suffer reduced fitness (Cramer *et al.,* 2016; reviewed in: Coyne and Orr, 2004). If individuals also discriminate against or preferentially mate with specific members of their own species, however, this non-random mating can inhibit gene flow within the species and begin the processes that eventually lead to the development of genetic isolating barriers between groups that were once considered a single species (Morris and Lundberg, 2011).

In *Drosophila*, cuticular hydrocarbons (CHCs) are key mediators of mate choice and recognition during courtship (reviewed in: Bontonou and Wicker-Thomas, 2014). The profile of CHCs that are expressed by an individual is dependent upon both its genetic background and the environment in which it lives and develops. Dramatic changes in CHCs due to the interaction of genotype with the environment can take effect within a single generation and affect large numbers of the population if they share similar environmental conditions, whereas genetic changes that alter CHC production are typically slow to reach high levels of prevalence in a population, a process which may take many generations (reviewed in: Ingleby, 2015). This means that individuals of the same species that develop in different environments may display dramatically different CHC profiles, despite their genetic compatibility. Since CHCs are intrinsically involved in the recognition of mates in *Drosophila,* it is possible that changes in CHC profiles result due to differences in environmental conditions and can lead to prezygotic isolating barriers. Here, I sought to determine how environmental changes related to the diet and gut microbiota impact the CHC profile of *D. melanogaster* flies.

3.1.1 Assortative mating in *Drosophila*

Within the same species, drosophilid flies can display high variability in the amounts of CHCs that they produce (Ingleby *et al.,* 2013). Flies with identical genotypes can express dramatically different CHC profiles if raised in different environments. Differences in diet can lead to assortative mating based on the ancestral diet type, and this preference is lost if flies are treated with antibiotics, indicating that the gut microbiota is responsible for this preference (Dodd, 1989; Sharon *et al.,* 2010). It is possible that since CHCs are so intimately involved in mate choice and recognition, they are the signals that are being altered by microbial content that allows flies to mate assortatively. This theory is supported by the evidence that CHCs change due to both diet type and antibiotic treatment that eliminates all commensal microbes (Sharon *et al., 2010*), although limited compounds were tested and a direct link between CHC variation and assortative mating was not tested. The selective assortative mating of flies raised in similar environments, and the response of CHCs to antibiotic treatment, indicates that the different microbial species maintained by flies raised on different food sources may be able to affect both preference for and expression of particular CHCs, if CHCs are indeed the mechanism by which flies are mating assortatively. It is not clear, however, precisely how microbial species impact the CHC profile expressed by flies independently of diet, and which compounds in the profile are affected. In order to uncover how microbes are acting to alter *Drosophila* mating preferences and to determine if this preference is mediated by CHCs, it is necessary to first understand how microbial content, diet, and the interaction of the two are able to affect the range of CHCs produced by flies.

3.1.2 Dietary effects on CHCs

Diet has been found to have a profound effect on the abundances of individual compounds and on the total amounts of CHCs produced in *D. melanogaster* flies. Fedina *et al.* (2012) tested the CHC profiles of female flies raised on high or low concentrations of yeast and sugar and discovered that they had opposite effects on CHC biosynthesis: high yeast concentrations led to a lower proportion of long-chain CHCs (which represents a more attractive profile (Kuo *et al.,* 2012)) and less attractive short-chain profiles, while high

sugar concentrations yielded the opposite results (a less attractive long-chain profile, represented by a greater proportion of long-chain CHCs, and a more attractive short-chain profile). Despite these changes in CHCs, the authors found that neither diet type altered female attractiveness to males significantly, and concluded that the diets tested led to overall neutral changes in CHC profiles, where some attractant compounds increased in prevalence while others decreased within each diet type, leading to no net change in attractiveness of the flies. Of particular note is that total abundance of CHCs was found to be similar regardless of diet type early in life, but later in life flies fed diets with higher yeast contents had dramatically increased (nearly two-fold) total CHC levels, regardless of the sugar content of their food. Yeast mainly provides flies with protein, but is also a source of lipids, fatty acids, and various vitamins (Fedina *et al.,* 2012). Increased dietary yeast could be causing increased CHCs by increasing total CHC production, increasing hydrocarbon transport to the cuticle, or some combination of the two.

3.1.3 Insulin signaling

One of the ways in which diet acts to impact CHC production is through insulin/insulin-like growth factor signaling (IIS) (Kuo *et al.,* 2012). Insulin/insulin-like growth factor signaling signaling is known to be intimately involved in the aging process, with greater levels of IIS leading to shorter-lived flies (Clancy *et al.,* 2001; Kuo *et al.,* 2012). Insulin/insulin-like growth factor signaling levels are tied to protein levels, with lower protein levels resulting in less IIS and longer-lived flies that age more slowly than flies fed high protein content (Fedina *et al.,* 2012). When IIS is knocked down in *Drosophila*, expression of genes involved in CHC biosynthesis including *eloF, Desat1, Desat2* and *DesatF* increases. Female flies that have IIS knocked down are longer-lived but less attractive to males, and display higher proportions of long-chain CHCs (representative of a less-attractive CHC profile) (Kuo *et al.,* 2012). When IIS is genetically increased, on the other hand, expression of *eloF* and *Desat1* decreases, flies are more attractive, and lifespan is reduced (Clancy *et al.,* 2001; Kuo *et al.,* 2012). Flies fed diets rich in protein (like those raised on high yeast diets) that experience high levels of IIS age more quickly than their low-protein counterparts, but are generally more attractive than flies with low IIS (Kuo *et al.,* 2012). Increased dietary protein therefore has a positive effect on the attractiveness of

flies and acts to decrease their lifespan, a process which is mediated by increased IIS in response to increased protein levels.

3.1.4 Commensal microbial species in *Drosophila*

The *Drosophila* gut microbiome typically contains some 30 or more distinct microbial species, dominated by species of the genera *Lactobacillus* and *Acetobacter* (Wong *et al.,* 2011). The microbial species present and their relative abundances within the *Drosophila* gut are mainly based upon the type of food the flies are reared on, and is not closely linked to the host *Drosophila* species (Staubach *et al.,* 2013). Microbial species are passed linearly from parents, primarily the mother, to offspring. When the female lays an egg, she defecates on it; when larvae emerge from the egg and eat the chorion (eggshell) and surrounding food, they also consume the feces, allowing microbial species contained within to take up residence in their guts (reviewed in: Dillon and Dillon, 2004).

3.1.5 Impact of commensals on *Drosophila* phenotype

In addition to their effect on CHCs, gut microbes have been shown to affect other aspects of *Drosophila* fitness. Flies that have had their commensal microbes eliminated are able to survive, however they undergo slowed development, likely due to reduced levels of IIS (Shin *et al.,* 2011; Newell and Douglas, 2014). Neither the adult body size of axenic (germ-free) flies nor their fecundity has been found to differ from that of conventionally raised flies, however they do show a reduced metabolic rate (Ridley *et al.,* 2012). Axenic flies also display elevated glucose and triglyceride levels (Newell and Douglas, 2014). When species normally found in the guts of drosopholid flies from the genera *Acetobacter* and *Lactobacillus* are reintroduced either individually or together to axenic flies, only those flies with species from both genera together displayed triglyceride levels that were similar to those of conventionally raised flies (Newell and Douglas, 2014), indicating that these two genera alone were able to recapitulate the effect on lipid metabolism of the entire normal microbial complement. The authors showed that only *Acetobacter* species were associated with a reduction in triglyceride levels towards normal levels, and that *Lactobacillus* species helped to promote *Acetobacter* abundance, rather than directly

affecting triglyceride levels. Interestingly, axenic flies that were found to have reduced levels of IIS recovered normal development in food supplemented with acetic acid, a byproduct of *Acetobacter* species metabolism (Shin *et al.,* 2011). It is possible that one of the ways that commensals influence the host is by modulating IIS signaling levels, and *Acetobacter* species in particular seem to be important in influencing the fly's nutritional phenotype. Although altered triglyceride levels are indicative of overall changes in fatty acid metabolism which may affect CHC levels, no data exists for how individual microbial species impact the CHC profile of flies, and how diet may modulate this interaction.

3.1.6 Environmental influences on CHCs

The underlying genetic complement is ultimately responsible for determining the types of compounds that may be produced by a fly, but the total amounts and ratios of each CHC will be dependent upon its interaction with the environment, both in the larval stage and during its adult life. Both diet and microbial content have been shown to impact the CHC profiles and attractiveness of flies (Sharon *et al.,* 2010; Fedina *et al.,* 2012; Kuo *et al.,* 2012); however, since the gut microbiota itself modulates how flies process their food source and likewise food source dictates the types and amounts of different microbial species that may be supported by the host fly, it is necessary to separate how gut microbial content and diet independently affect the CHC profiles of flies, and how their interaction may give rise to the distinct CHC profiles that may form the basis for mating bias. I therefore sought to independently test how each factor influences CHCs in female *D. melanogaster* flies by altering microbial composition across two different food types, one in which protein is replete and IIS is presumed to be high, and one in which protein concentration is lower and IIS should be reduced. Flies that were raised conventionally, axenic flies, and flies to which individual species of (presumed) *Acetobacter* or *Lactobacillus* were added were tested and their CHCs profiles compared.

3.2 Methods

3.2.1 Maintenance of *Drosophila* stocks

Stocks were maintained in 30 mL vials at 24°C on a 14:10 hour light:dark cycle at 76% relative humidity. *Drosophila melanogaster* line Canton-S (obtained from Dr. Anne Simon) was used. *Drosophila* stocks were maintained on a standard cornmeal-yeast diet (CMY) (Bloomington Drosophila Stock Center standard recipe).

3.2.2 Diet and microbial treatment groups

When assayed, *Drosophila* were either maintained on the standard CMY diet or were transferred to a modified diet that was identical but for a reduction by half of the yeast concentration compared to the standard diet (half-yeast, HY). Flies raised on HY food should have lowered levels of IIS, due to reduced protein content in their diets (Fedina *et al.,* 2012). Within each of the two diet types, one of six treatment types was used to vary the microbial composition of the flies (12 treatments in total): i) a control stock of conventionally raised flies; ii) axenic flies; iii) axenic flies to which microbes had been reintroduced from control fly (i) feces; iv) axenic flies to which a (presumed; see section 3.25 below) *Acetobacter* species has been added; v) axenic flies to which a (presumed; see section 3.2.5 below) *Lactobacillus* species has been added, and vi) axenic flies to which both presumed *Acetobacter* and *Lactobacillus* species have been added. Flies were added to their respective treatment vials following mild CO_2 anesthesia on CO_2 pads that were covered by a fresh kimwipe each time to avoid contamination from the CO_2 pads themselves.

3.2.3 Generation of axenic flies via antibiotic treatment

Axenic flies were generated by cultivating flies on a diet (either CMY or HY) supplemented with 50 µg/mL tetracycline, 200 µg/mL rifampicin and 100 µg/mL streptomycin, as per Sharon *et al* (2010). Flies were confirmed to be axenic by crushing and streaking on MRS (10 g/L peptone; 8 g/L meat extract; 4 g/L yeast extract; 20 g/L

glucose; 2 g/L dipotassium hydrogen phosphate; 5 g/L sodium acetate trihydrate; 2 g/L triammonium citrate; 0.2 g/L magnesium sulfate heptahydrate; 0.05 g/L manganous sulfate tetrahydrate; 1.5% agar) and BHI (brain-heart infusion, porcine, 1.5% agar) agar plates and incubating at 30°C. Flies treated with antibiotic did not produce colonies on either type of plate after 4 days of incubation, confirming the absence of any species present in conventionally reared flies that grow under these conditions, whereas control flies streaked onto these plates display isolated colonies after overnight growth and widespread growth over the entirety of the plates after 4 days. Axenic flies may still harbor varieties of bacteria that do not grow well on either MRS or BHI media or that require strictly anaerobic conditions for growth. Flies that developed from eggs that were laid on the antibiotic-supplemented food (rather than adults that were transferred from regular food) were used in all cases.

3.2.4 Reintroduction of normal microbial content to axenic flies

Male or virgin female flies that had been raised in typical lab conditions were transferred to CMY or HY food after collection under light CO_2 anesthesia and allowed to live for a period of 5 days in order to transfer their feces (containing their gut microbes) to the food. Flies were then removed, and newly-eclosed (0-8 hours old, collected under light CO_2 anesthesia on a fresh kimwipe) axenic flies were added to the vials. To confirm that the offspring of these axenic flies were able to take up microbes from the media via fecal ingestion, 5 day old adult offspring of the axenic flies were homogenized in sterile DI water and streaked on MRS and BHI agar plates, and colony growth was observed. Five flies per treatment were homogenized and streaked to confirm the absence or presence of microbial species.

3.2.5 Isolation and cultivation of *Drosophila* gut microbial species

Flies grown under normal conditions (five total, pooled) were crushed via mortar and pestle in 1.5 mL of DI water in a sterile Eppendorf tube and vortexed. A 100 µL aliquot of the solution was streaked onto either MRS or BHI agar plates. MRS media is used for the cultivation of *Lactobacillus* species (De Man *et al.,* 1960) and BHI can be used to grow

Acetobacter species (Sharafi *et al.,* 2010); these species represent the most abundant bacterial families found in *Drosophila* (Wong *et al.,* 2011). Colonies on MRS and BHI plates all had similar morphology both within and between plates (smooth, white colonies). Both *Lactobacillus* and *Acetobacter* species are known to form colonies with this morphology; although the different media types were selected for optimal growth of their target bacterial species, it is possible that species isolated on both media types were the same species. PCR using general 16S rRNA primers (forward: 16SA1: 5'-AGAGTTTGATCMTGGCTCAG-3'; reverse: 16SB1: 5' – TACG GTACCTTGTTACGACTT-3') was performed as per Ridley *et al.* (2012) in order to identify the genus of the species isolated, but was unsuccessful using these primers. Individual colonies from each MRS (putative *Lactobacillus* species) and BHI (putative *Acetobacter* species) were then used to inoculate MRS or BHI liquid broth, respectively, and were grown with shaking at 30°C. Glycerol stocks of both species were generated and stored at -80°C for future testing or species determination.

3.2.6 Introduction of isolated microbial species to axenic flies

Individual colonies from plates streaked with presumed microbial species *Acetobacter* or *Lactobacillus* were grown overnight in liquid culture. Cells were then centrifuged and re-suspended in sterile DI water to a final concentration of 10^8 CFU/mL. Three separate bacterial treatments were added to the surface of either HY or CMY food, as per Newell and Douglas (2010): 50 µL of the re-suspended cells of presumed species *Acetobacter* or *Lactobacillus*, or 25 µL of each to the same vial. Newly-eclosed axenic flies were then added to the treated vials and allowed to reproduce.

3.2.7 Collection of flies and extraction of CHC content

Freshly-eclosed flies (0-8 hours of age) from each of the treatment types were collected via mild CO_2 anaesthesia 0-12 hours after eclosing. Flies were then separated based on sex and females were transferred to collection vials that had been treated in the same way as the vials they had been laid in (both in terms of diet type and microbial treatment). Female flies were aged to 5 days and CHCs were extracted 2 hours after lights-on as previously

described (see section 2.2.5). Five flies from each treatment type (12 treatments total; 6 per diet type) were analyzed.

3.2.8 Analysis of CHC content of flies

Gas chromatography conditions and analysis of GC output was performed as previously described (see section 2.2.6) to determine the total peak area of each compound that was detected and to adjust this area based on the area of the internal standard peak. Fourteen compounds were compared, as these are the compounds that were most consistently identified using this method. The average mass of flies from each of the separate treatments was used to adjust the GC output to control for possible variation in CHC abundance due to the body size of the flies by multiplying the peak areas obtained from the GC output by a ratio between the average mass of control flies for that diet type and the average mass of flies from the relevant treatment type. Data were analyzed using both one-way (comparing within diet type to look for differences in CHCs due to microbial species) and two-way ANOVA (comparing within and between diet type to look for differences due to diet, microbial species, or their interaction) to compare the mean peak area of each compound and determine if any compounds differ significantly across or within the different treatment types. As previously described, the relative abundance of compounds was also compared (see section 2.2.6). A false discovery rate (FDR) correction was used to account for multiple tests on the same data (3 ANOVAs and 5 comparisons per compound, 14 compounds).

3.3 Results

3.3.1 Impact of gut microbiota on CHCs

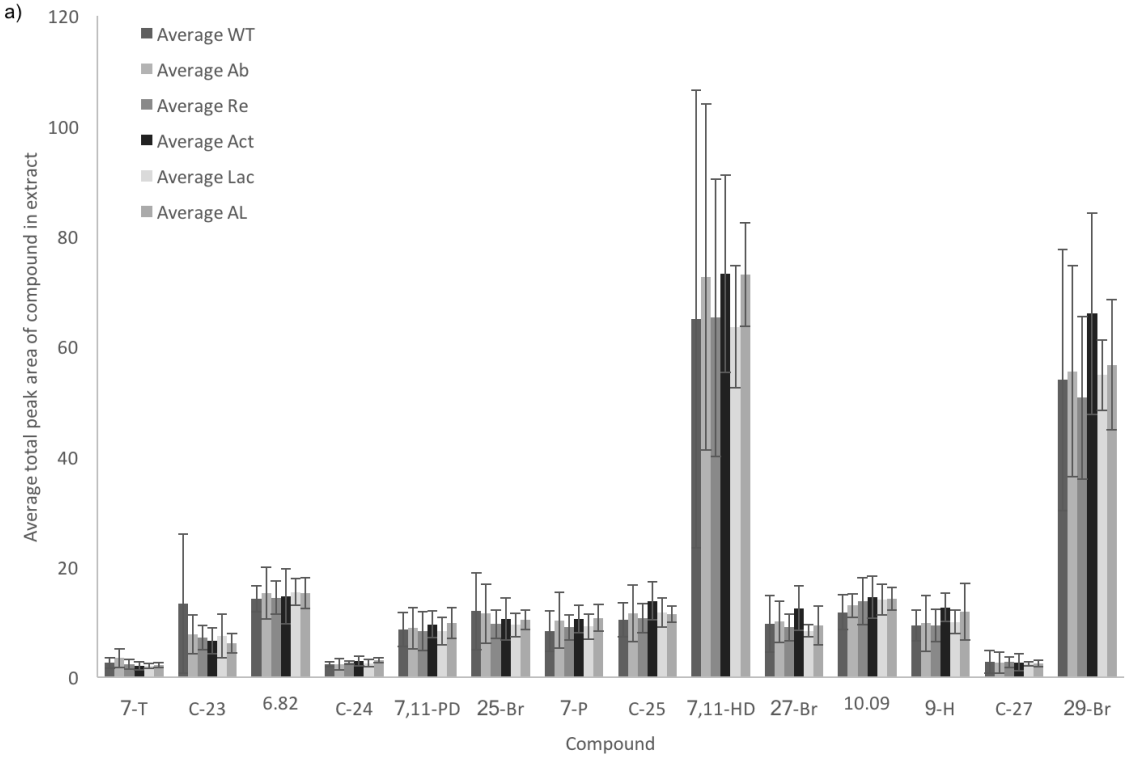
Microbial content was not found to have a significant effect on any of the CHCs measured, or on total CHC abundance (Figure 3.1; see appendix B). Axenic flies did not show significantly different CHC levels to control flies or to flies to which conventional microbes had been reintroduced, or to gnotobiotic flies to which individual microbial species had been introduced.

3.3.2 Impact of diet on CHCs

Similar to the results found in Chapter 2, significance or non-significance of relative abundances mirrored the results for total peak areas of compounds; in all cases where a significant change in the total peak area of a compound was observed, this compound was also found to significantly differ in relative abundance, and in no cases did a compound differ significantly in relative abundance and not in total peak area.

Across all treatment types, diet (CMY vs HY) was found to have a significant effect on the abundance of 7,11-HD, 2-methylhexacosane (27-Br), *n*-heptacosane (C-27), and 2-methyloctacosane (29-Br) ($p<0.0002$, ANOVA, $n=5$ for all compounds). Flies fed a half-yeast diet had significantly higher levels of each of these compounds than flies fed a CMY diet, and this effect was independent of microbial content (Figures 3.2 and 3.3; see appendix B).

a)


Average WT
Average Ab
Average Re
Average Act
Average Lac
Average AL
Average total peak area of compound in extract
120
100
80
60
40
20
0
7-T
C-23
6.82
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
27-Br
10.09
9-H
C-27
29-Br
Compound

b)

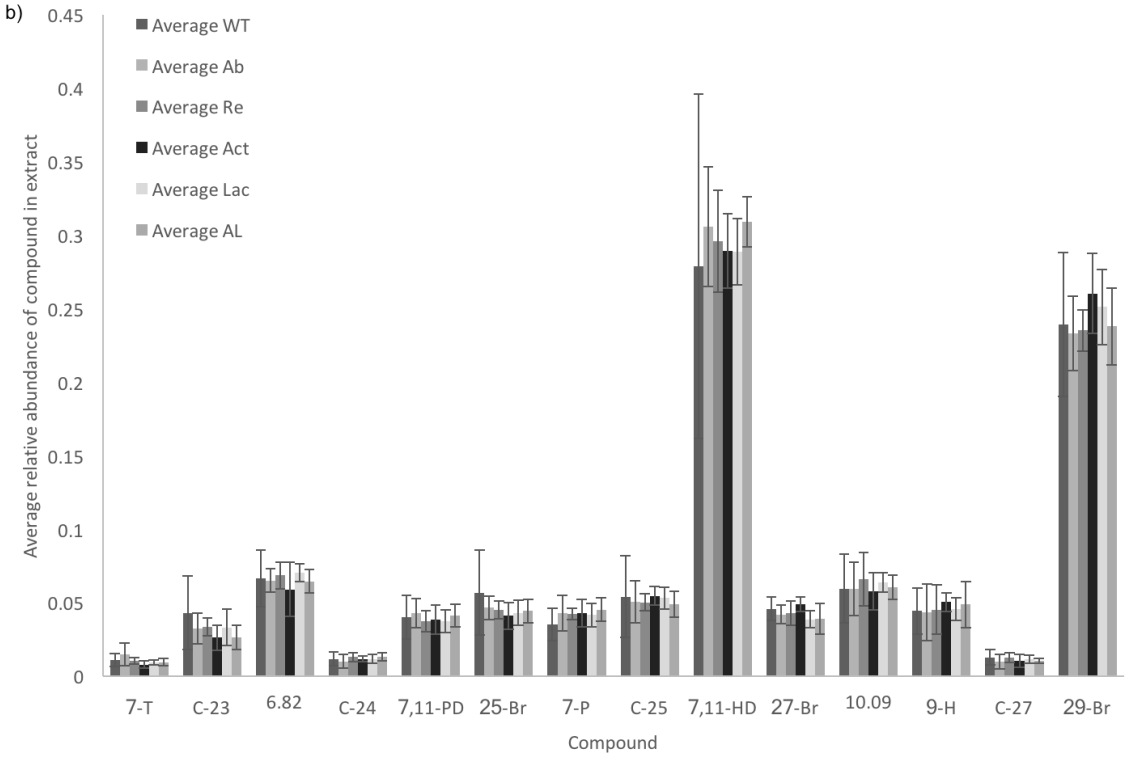
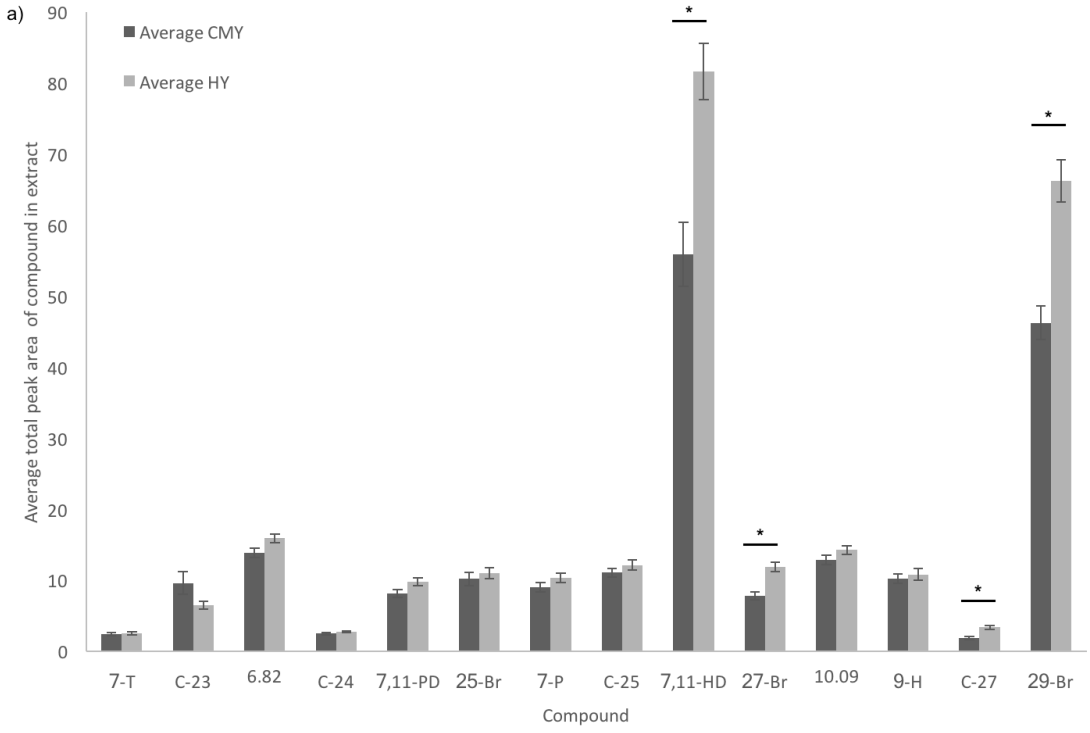

Average WT
Average Ab
Average Re
Average Act
Average Lac
Average AL
Average relative abundance of compound in extract
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
7-T
C-23
6.82
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
27-Br
10.09
9-H
C-27
29-Br
Compound

Figure 3.1: Effect of microbial treatment on CHCs in female *D. melanogaster* flies. Bars represent the mean (±SD) total peak area or relative abundance of compounds extracted from the cuticles of 5-day old females from each of the microbial treatments. WT = wild-type (control); Ab = antibiotic treated (axenic flies); Re = reseeded with control microbes; Act = presumed *Acetobacter* species added to axenic flies; Lac = presumed *Lactobacillus* species added to axenic flies; AL = both presumed *Acetobacter* and *Lactobacillus* species added to axenic flies. a) Average total peak area of compounds across the treatment types; b) average relative abundance of compounds across the treatment types.


a)
Average CMY
Average HY
Average total peak area of compound in extract
0
10
20
30
40
50
60
70
80
90
*
*
*
*
7-T
C-23
6.82
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
27-Br
10.09
9-H
C-27
29-Br
Compound

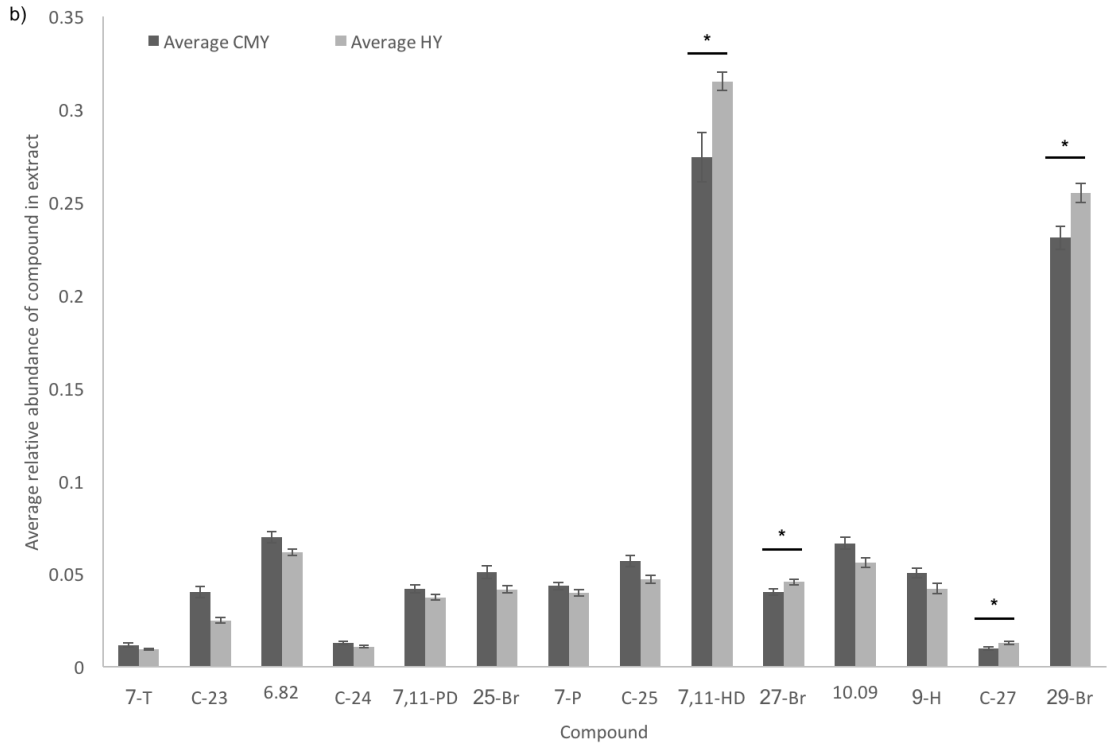

b)
Average CMY
Average HY
Average relative abundance of compound in extract
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
*
*
*
*
7-T
C-23
6.82
C-24
7,11-PD
25-Br
7-P
C-25
7,11-HD
27-Br
10.09
9-H
C-27
29-Br
Compound

Figure 3.2: Effect of diet type on CHCs in female *D. melanogaster* flies. Bars represent the mean (±SD) total peak area or relative abundance of compounds extracted from the cuticles of 5-day old females from each of the diet types (CMY or HY). a) Average total peak area of compounds across the diets; b) average relative abundance of compounds across the diets.

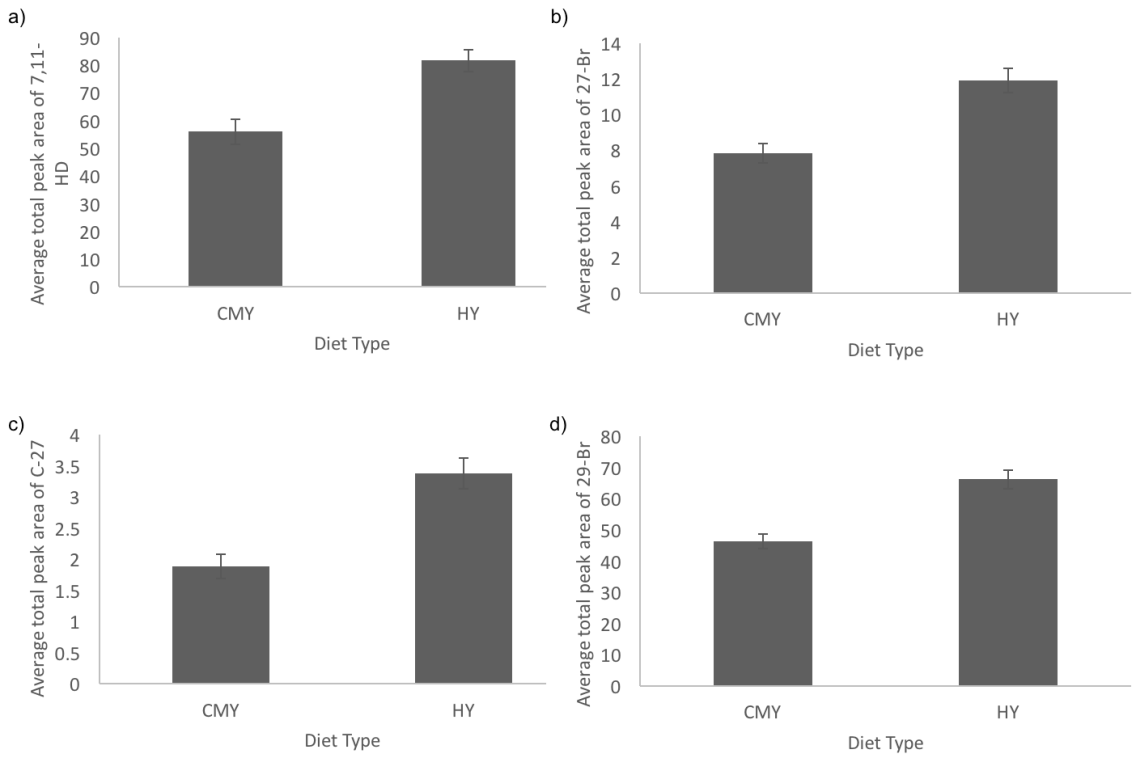


Figure 3.3: Average total peak area of specific CHCs in flies fed a CMY or a HY diet. Bars represent the mean (±SD) total peak area or relative abundance of compounds extracted from the cuticles of 5-day old females from each of the diet types (CMY or HY). a) Average peak area of 7,11-HD (p_{diet}=1.530x10^{-4}, ANOVA, t=-7.208, d.f.=29); b) average peak area of 27-Br (p_{diet}=1.698x10^{-4}, ANOVA, t=6.800, d.f.=29); c) average peak area of C-27 (p_{diet}=1.538x10^{-4}, ANOVA, t=-6.737, d.f.=29); d) average peak area of 29-Br (p_{diet}=4.458x10^{-4}, ANOVA, t=-7.529, d.f.=29). n=30 for each diet type.

3.4 Discussion

The environment in which a fly is reared may have a direct effect on its attractiveness as a mate (Kuo *et al.,* 2012). This effect may be mediated in part by differential expression of CHCs depending on the diet and microbial complement present in the flies. Since mate preference can change with microbial content (Sharon *et al.,* 2010) and CHC profiles represent a key component in mate choice, discovering how exactly CHCs change with microbial content is the first step in determining whether or not they may represent a potential mechanism by which commensals are acting to affect mate choice. Here, I looked at how microbial species isolated from the guts of laboratory-reared *D. melanogaster* impact the CHC profiles of female flies, and if this effect could be modulated by diet.

Despite previous results that found that flies treated with antibiotics showed a dramatic decrease in several of their most abundant CHC compounds when compared to conventionally reared flies (Sharon *et al.,* 2010), the CHCs of flies in this study remained unchanged across all microbial treatment types, including antibiotic treatment. Flies raised conventionally had CHC profiles that were indistinguishable from axenic flies, or from those flies to which presumed microbial species had been reintroduced. It is possible that antibiotic treatment, while successful in removing aerobic species as evidenced by the lack of colony growth on MRS or BHI media following treatment, was unsuccessful in removing all microbial species from the flies, although a study using the same methods to generate axenic flies did find changes in CHCs (Sharon *et al.,* 2010). Possibly certain anaerobics or species that are difficult to cultivate in a lab setting persist within the flies, and these species may be the ones responsible for the conventional CHC profile, while their removal results in the shifts in CHC abundance in response to antibiotic treatment that were previously noted by Sharon *et al* (2010). Alternatively, as was noted in chapter 2, only a small subset of the total CHC complement was able to be studied here; Sharon *et al.* found that microbial complement impacts some of those compounds that were not able to be detected in this study, but those compounds that were detected both here and by Sharon *et al.* did not show the same trends due to microbial treatment. Finally, it is possible that microbial species may only have a large impact on CHCs when flies are fed diets that are

very low in nutritive content, as was studied in Sharon *et al.* (2010). In order to more reliably determine whether or not microbial content may affect CHCs, future work in which axenic flies are generated in a different manner (e.g. via dechorionation (Ridley *et al.,* 2012)) or in which flies are raised on a far less nutritious diet and microbial effects tested should be done to confirm whether the CHC profiles of truly axenic flies are markedly different from those of conventionally reared flies, and whether microbes are perhaps more important and have greater effect on CHCs in less nutritious circumstances. If it is true that CHC profiles remain unchanged regardless of microbial complement, even between conventional and axenic flies, another explanation must be found for the assortative mating due to gut microbial complement that was observed in the study by Sharon *et al.* (2010). For example, it is possible that microbes are acting to alter perception and behaviour at some other step of the courtship ritual, such as during the male's wing-song, rather than influencing the biosynthesis of chemical signals in the flies.

Despite conflicting results when microbial species were altered across different treatments, flies raised on different diets showed clear trends in CHC changes across the different diet types. Flies fed a diet that was low in yeast had increased levels of long-chain CHCs, with no significant changes in any short-chain compounds. In particular, HY flies had significantly increased amounts of long-chain CHCs 7,11-HD, 27-Br, C-27, and 29-Br. This is consistent with previous studies that found that higher concentrations of dietary yeast led to a more attractive long-chain CHC profile that had lower levels of long-chain CHCs (Fedina *et al.,* 2012). Reducing the yeast concentration below balanced levels was found here to have the opposite effect on long-chain CHCs, leading to larger amounts of these compounds and a presumably less attractive profile. Although 7,11-HD is a known attractant molecule (Coyne *et al.*, 1994), no role in mating for 27-Br, C-27 or 29-Br has been suggested, although a higher proportion of long-chain CHCs generally is known to represent a less attractive profile (Kuo *et al.,* 2012). Flies that have been fed diets that have reduced yeast concentrations have lower levels of dietary protein, and therefore presumably have reduced levels of IIS (Fedina *et al.,* 2012). As shown by Kuo *et al.* (2012), this results in increased expression of CHC biosynthetic genes such as *eloF, Desat1, Desat2* and *DesatF*, which explains the increased proportion of long-chain CHCs being produced by these flies with low levels of IIS.

Although 7,11-HD is conventionally thought of as an attractant in *Drosophila* and it is known to stimulate courtship behaviours such as wing vibration (Antony and Jallon, 1982), an increase in this compound alongside other long-chain CHCs has been shown to represent a less attractive CHC profile (Kuo *et al.,* 2012). This may be explained by the suggestion that the major role of 7,11-HD is to negate or cover up the repellent effect of the cVA that is transferred from males to females and marks them as mated rather than to act as a general attractant, as the application of 7,11-HD to oenocyteless flies that are unable to produce any CHCs was not found to result in a decrease in mating latency, calling into question its role as an aphrodisiac (Billeter *et al.,* 2009). Although higher levels of 7,11-HD have not been found to have a repellent effect on *D. melanogaster* males, it seems likely that an increase in this compound is neutral besides the context in which a female has already been mated once, and is not enough to limit the detrimental effect of an overall longer-chain CHC profile on attractiveness.

Total amounts of CHCs did not differ significantly between the diet types, indicating that some compounds in HY flies must have decreased in abundance with the concomitant and significant increase in certain long-chain compounds, though not enough to reach the significance threshold. It is possible that many compounds saw a slight decrease in relative abundance, or that some of the compounds that could not be detected here were altered. This is consistent with previous studies that have found that diet only impacts total CHC amounts later in life (day 23+), and that flies fed widely varying diets produced very similar amounts of total CHCs early in life (Fedina *et al.,* 2012).

Diet has a clear effect on CHC profiles in *D. melanogaster*, with the role of microbes in CHC biosynthesis somewhat more clouded due to contradictory results. Flies that are raised in less nutritious circumstances, here represented by the HY condition, display CHC profiles that are consistent with overall less attractive profiles (Kuo *et al.,* 2012) than those of flies raised in conditions where protein is replete. This makes sense in the context of CHCs as chemical signals used in mate choice, as an honest signal of the nutritional state of a fly is important in the recognition of quality mates (reviewed in: Bontonou and Wicker-Thomas, 2014; Howard and Blomquist, 2005; Kuo *et al.,* 2012). More work must be done in order to determine whether microbial content can influence CHC levels, and if so,

whether these changes can be sensed by flies raised in a similar environment and lead to assortative mating.

CHC profiles are complex, with the relative abundances of certain compounds as well as their absolute amounts factoring into the perception of that profile; it would be interesting to discover if microbial species are able to affect CHC profiles so that they give accurate information not only about the overall health of the fly, but also about which commensals it was reared with, without interfering with (or possibly *by* its interference with) the honest signals of the fly's nutritional state. It is possible that diet and microbial content have distinct roles in mate choice, where diet acts to modulate CHCs to give signals pertaining to the overall health of the fly, and microbial content acts in some currently unknown way to signal a common environmental background that leads to assortative mating. If diet alters CHCs in a predictable way wherein flies from highly nutritious environments are more attractive, these flies should always be more attractive and be preferentially mated with, regardless of the diet the choosing flies were raised on. It is only when microbes are introduced that flies begin to mate assortatively with individuals raised in similar conditions; it is therefore necessary that microbial content must be able to alter some trait(s) and/or the preference for that trait, rather than simply acting to make flies more healthy generally (in which case the healthiest flies would be preferentially chosen, rather than those from the same environment). Future work examining how perception changes in flies with different microbial complements is needed to elucidate how flies are discriminating against those that were raised in different environments.

3.5 Bibliography

Antony, C. and Jallon, J-M. **1982**. The chemical basis for sex recognition in *Drosophila melanogaster*. *Journal of Insect Physiology*. **28**:873-880.

Billeter J.C., Atallah J., Krupp J.J., Millar J.G. and Levine, J.D. **2009**. Specialized cells tag sexual and species identity in *Drosophila melanogaster*. *Nature*. **461**:987–991.

Bontonou, G. and Wicker-Thomas, C. **2014.** Sexual communication in the *Drosophila* genus. *Insects*. **5**:439-4583.

Clancy, D., Gems, D., Harshman, L.G., Oldham, S. and Stocker, H. **2001**. Extension of life-span by CHICO, a *Drosophila* insulin receptor substrate protein. *Science*. **292**:104-106.

Coyne, J., Mah, K. and Crittenden, A. **1994**. Genetics of a pheromonal difference contributing to reproductive isolation in *Drosophila*. *Science*. **265**:1461-1464.

Coyne, J.A. and Orr, H.A. **2004**. Speciation. Sunderland, Mass: Sinauer Associates.

Cramer, E.R., Ålund, M., McFarlane, E., Johnsen, A. and Qvarnström, A. **2016**. Females discriminate against heterospecific sperm in a natural hybrid zone. *Evolution*. **70**:1844-1855.

De Man, J.C., Rogosa, M. and Sharpe, M. **1960**. A medium for the cultivation of *Lactobacilli*. *Journal of Applied Bacteriology*. **23**:130-135

Dillon, R.J. and Dillon, V.M. **2004**. The gut bacteria of insects: nonpathogenic interactions. *Annual Review of Entomology*. **49**:71-92.

Dodd, D.M. **1989**. Reproductive isolation as a consequence of adaptive divergence in *Drosophila pseudoobscura*. *Evolution*. **43**:1308-1311.

Fedina, T.Y., Kuo, T-H., Dreisewerd, K., Dierick, H.A., Yew, J.Y. and Pletcher, S.D. **2012**. Dietary effects on cuticular hydrocarbons and sexual attractiveness in *Drosophila*. *Public Library of Science*. **7**:e49799.

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

Ingleby, F.C., Hosken, D.J., Flowers, K., Hakwes, M.F., Lane, S.M., Rapkin, J., Dworkin, I. and Hunt, J. **2013**. Genotype-by-environment interactions for cuticular hydrocarbon expression in *Drosophila simulans*. *Journal of Evolutionary Biology*. **26**:94-107.

Ingleby, F.C. **2015**. Insect cuticular hydrocarbons as dynamic traits in sexual communication. *Insects*. **6**:732-742.

Kuo, T-H., Fedina, T.Y., Hansen, I., Dreisewerd, K., Dierick, H.A., Yew, J.Y. and Pletcher, S.D. **2012**. Insulin signaling mediates sexual attractiveness in *Drosophila*. *Public Library of Science*. **8**:e1002684.

Morris, D.W. and Lundberg, P. **2011**. Pillars of Evolution: Fundamental Principles of the Eco-evolutionary Process. New York: Oxford University Press.

Newell, P.D. and Douglas, A.E. **2014**. Interspecies interactions determine the impact of the gut microbiota on nutrient allocation in *Drosophila melanogaster*. *Applied and Environmental Microbiology*. **80**:768-796.

Ridley, E.V., Wong, A.C-N., Westmiller, S. and Douglas, A.E. **2012**. Impact of the resident microbiota on the nutritional phenotype of *Drosophila melanogaster*. *Public Library of Science*. **7**:e36765.

Sharafi, S.M., Rasooli, I. and Beheshti-Maal, K. **2010**. Isolation, characterization and optimization of indigenous acetic acid bacteria and evaluation of their preservation methods. *Iranian Journal of Microbiology*. **2**:38-45.

Sharon, G., Segal, D., Ringo, J.M., Hefetz, A., Zilber-Rosenberg, I. and Rosenberg, E. **2010**. Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA*. **107**:20051-20056.

Shin, S.C., Kim, S-H., You, H., Kim, B., Kim, A.C., Lee, K-A., Yoon, J-H., Ryu, J-H. and Lee, W-J. **2011**. *Drosophila* microbiome modulates host development and metabolic homeostasis via insulin signaling. *Science*. **334**:670-674.

Staubach, F., Baines, J.F., Künzel, S., Bik, E.M. and Petrov, D.A. **2013**. Host species and environmental effects on bacterial communities associated with *Drosophila* in the laboratory and in the natural environment. *Public Library of Science*. **8**:e70749.

Wong, C.N., Ng, P. and Douglas, A.E. **2011**. Low-diversity bacterial community in the gut of the fruitfly *Drosophila melanogaster*. *Environmental Microbiology*. **13**:1889-1900.

Chapter 4

4 Conclusions and future work

In studying speciation, there are two major areas of inquiry that can be used to lead to the refinement of what we consider to be a species: the concept of how one group can, over time, split into two and how, once they have split, these groups can be maintained as distinct entities and avoid collapsing into a single group that experiences free gene flow. If we can understand how and why groups differentiate over time, we can, perhaps, come closer to understanding the parameters that define a species and hopefully come to a concrete and universal resolution as to what a species really is.

In sexually reproducing species, mate choice and sexual selection are key factors in determining the course of evolution, as it will determine whether or not gene flow may occur between genetically compatible groups and is one of the forces that may ultimately lead to the development of genetic, as well as behavioural, incompatibilities (reviewed in: Jones and Ratterman, 2009; Carson, 2003). In *Drosophila*, chemical communication is of primary import in mate choice (reviewed in: Stieger and Stökl, 2014). The expression of different CHC profiles is the result of the interaction of the genetic background of an individual with the environment in which it was raised (Foley *et al.,* 2007; reviewed in: Ingleby, 2015), and has been found to constitute a prezygotic reproductive barrier in *Drosophila* (reviewed in: Bontonou and Wicker-Thomas, 2014). I therefore sought to discover precisely how environmental changes can influence the CHC profile expressed by flies of the same currently accepted species, and what genetic factors are present between species that have already diverged into distinct groups that may influence their CHC profiles. I discovered that a *D. melanogaster* gene, *CG5946*, and its *D. simulans* homolog *GD12773* lead to differences in the abundance of sex pheromones between these species and their hybrids, potentially through differential expression. I also confirmed that within the species *D. melanogaster,* flies that are raised on different diet types show altered CHC profiles, although gut microbial content was not found to have an effect on the CHCs of these flies.

While the environment may influence CHC expression rapidly within a single generation, genetic changes that can influence CHCs represent more stable prezygotic barriers that may be maintained over many generations. If species are to diverge in sympatry, it is necessary that flies first begin to mate non-randomly within the population, and that this non-random mating becomes genetically fixed in order to stably prevent the gene flow that impairs the development of reproductive isolating barriers (reviewed in: Coyne and Orr, 2004). Since CHCs are intrinsically involved in mate choice in *Drosophila*, it seems plausible that environmentally-caused changes to CHCs may be some of the initial factors that, if altered, could affect mate choice and cause flies to begin mating non-randomly. Studying how CHCs can be affected within a generation within a single species may therefore help to answer questions about how single groups may initially split into independent populations, whereas studying the existing genetic variation giving rise to different CHC profiles across species can yield insights into how these initial transient changes might become stably integrated into the genome over the course of evolutionary time.

Although this study has uncovered some of the factors that may affect *Drosophila* CHC profiles, a link must be established between the altered CHCs and the mating preferences and success of these flies in order to establish that these alterations can influence mating propensity within and between species. In particular, it would be beneficial to discover whether hybrids lacking the *D. melanogaster* copy of *CG5946*, which produce a more *simulans*-like CHC profile, are more or less attractive to *D. simulans* males than their pure *D. melanogaster* mothers or WT hybrids. It would also be interesting to discover where and when this gene is expressed in both *D. melanogaster* and *D. simulans*, and how its product may function outside of fatty acid metabolic processes. With respect to how environment may influence CHCs within a species, it remains to be determined whether the altered CHC profiles of flies, as the result of their diet, also impacts their attractiveness to males, and whether or not males from the same diet preferentially mate with females that share their diet. In other words, whether the same parameters that are able to alter CHC expression are also able to alter perception and preference; future studies should therefore seek to establish a link between the altered CHC profiles found here and their effects on attractiveness and mate choice in *Drosophila*.

4.1 Bibliography

Bontonou, G. and Wicker-Thomas, C. **2014.** Sexual communication in the *Drosophila* genus. *Insects*. **5**:439-4583.

Carson, H.L**. 2003**. Mate choice theory and the mode of selection in sexual populations. *Proceedings of the National Academy of Sciences USA*. **100**:6584-6587.

Coyne, J.A. and Orr, H.A. **2004**. Speciation. Sunderland, Mass: Sinauer Associates.

Foley, B., Chenoweth, S.F., Nuzhdin, S.V. and Blows, M.W. **2007**. Natural genetic variation in cuticular hydrocarbon expression in male and female *Drosophila melanogaster*. *Genetics*. **175**:1465-1477.

Ingleby, F.C. **2015**. Insect cuticular hydrocarbons as dynamic traits in sexual communication. *Insects*. **6**:732-742.

Jones, A.G. and Ratterman, N.L. **2009**. Mate choice and sexual selection: What have we learned since Darwin? *Proceedings of the National Academy of Sciences USA*. **106**:10001-10008.

Sharon, G., Segal, D., Ringo, J.M., Hefetz, A., Zilber-Rosenberg, I. and Rosenberg, E. **2010**. Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA*. **107**:20051-20056.

Stieger, S. and Stökl, J. **2014**. The role of sexual selection in the evolution of chemical signals in insects. *Insects*. **5**:423-438.

Appendices

Appendix A: Average total peak area and relative abundance of compounds in disruption lines

Table A1: Average total peak area of compounds for disruption of *CG5946*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-22	1.991007304	0.843988273	1.999238744	0.523591697	4.457694898	1.080722597	3.891153033	0.581299319	0.5459
23-Br	0.051042388	0.004559301	0.051796824	0.003101615	20.40383872	2.61474343	12.80134039	2.01837336	0.1454
7-T	**5.038981158**	**0.831492859**	**2.676822518**	**2.278044829**	**23.28659576**	**3.111528198**	**58.67263348**	**10.13743212**	**0.0003**
C-23	20.08020903	2.522961956	21.37260484	11.20117994	69.56069608	14.30861696	53.47067934	1.636044648	0.1408
C-24	2.620828756	0.516535365	2.28826755	0.636105742	3.141872407	2.874092274	4.764961808	2.737409971	0.4275
7,11-PD	15.70977621	3.012589629	1.862928946	3.134072862	56.98352089	13.39612177	47.20679678	5.454457227	0.6531
25-Br	0.051042388	0.004559301	0.051796824	0.003101615	15.30108803	1.599665131	13.22871219	3.808978717	0.4101
7-P	15.26141953	2.256842591	27.56874026	1.705174193	15.31026218	4.003030256	27.84019992	5.854694296	0.9610
C-25	10.53278103	1.11706281	16.57089771	2.066159479	28.413742	5.424206696	21.38403683	3.514509954	0.0190
7,11-HD	**48.54487746**	**8.737886841**	**55.89977009**	**5.509375647**	**163.2067694**	**10.57634947**	**86.32153143**	**10.57191232**	**4.258E-05**

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A2: Average relative abundance of compounds for disruption of *CG5946*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-22	0.016461807	0.005771679	0.01567372	0.005666091	0.011047094	0.001713559	0.011779212	0.000941477	0.7597
23-Br	0.00042733	2.28539E-05	0.00040002	4.60076E-05	0.051514496	0.00987941	0.03875351	0.003793918	0.0706
7-T	**0.041993179**	**0.003943705**	**0.019527298**	**0.016645624**	**0.058049228**	**0.002919074**	**0.177514363**	**0.02053708**	**1.65E-05**
C-23	0.167484214	0.001770949	0.159155235	0.074094688	0.172740996	0.020568036	0.16314962	0.013954926	0.9783
C-24	0.021779648	0.002304728	0.017971987	0.006856087	0.007528222	0.006683047	0.014049472	0.006601678	0.1697
7,11-PD[3]	**0.130386908**	**0.011952933**	**0.01399887**	**0.02352253**	**0.142800927**	**0.034811707**	**0.143283276**	**0.010166274**	**0.0019**
25-Br	0.00042733	2.28539E-05	0.00040002	4.60076E-05	0.038343348	0.004252996	0.039677585	0.007304482	0.7873
7-P	0.128352415	0.021581647	0.213278538	0.029472848	0.038397606	0.010619891	0.083906736	0.009109973	0.1190
C-25	0.089666006	0.021916615	0.127284698	0.011896759	0.070622333	0.007429339	0.065850031	0.016573528	0.0445
7,11-HD	**0.403021163**	**0.027760666**	**0.432309613**	**0.066816525**	**0.408955751**	**0.028600916**	**0.262036196**	**0.019964457**	**0.0052**

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.
[3]Entry passes FDR, but differences are not due to *Sim/Bal* vs *Sim/Dis* (due to *Mel/Dis*).

Table A3: Average total peak area of compounds for disruption of *CG1945*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	$p^{1,2}$
C-22	2.79837885	1.364635923	3.597406818	2.427633182	2.460684768	2.785597911	7.588926709	3.124758196	0.28784
23-Br	0.052268547	0.005492879	0.416206701	0.739706219	11.73968087	4.740209289	13.18477922	3.317996711	0.0574
7-T	6.48087717	2.153861567	6.967889893	2.022614174	91.24138836	24.23827367	102.5340594	38.58209436	0.0027
C-23	22.44098495	3.203358308	34.37785811	6.680661298	55.78537954	10.03282321	98.39459208	26.36499847	0.2960
C-24	2.769681424	0.644251335	3.395100098	0.768529557	2.117081396	2.68318584	8.542836004	2.984584576	0.0986
7,11-PD	8.733223506	1.327476498	11.36666637	2.802353152	48.64451453	18.8674347	61.00857181	17.62750909	0.1416
25-Br	0.052268547	0.005492879	10.40551996	1.852275437	17.93871797	13.50663072	25.1064418	11.62133759	0.1385
7-P	0.052268547	0.005492879	17.12887537	6.848535725	0.065520368	0.015930069	65.06440225	24.99018386	0.0187
C-25	10.40222188	0.594360462	16.05850891	1.879861503	18.95500859	3.784655016	41.33848907	11.63613461	0.0413
7,11-HD	62.60674353	6.620821071	71.41541197	16.6433741	109.4450236	10.49762477	126.4023154	42.53006479	0.6386

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A4: Average relative abundance of compounds for disruption of *CG1945*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-22	0.023353944	0.008949142	0.020290736	0.014177641	0.006042735	0.006783132	0.013498792	0.003258905	0.2746
23-Br	0.000452354	5.43984E-05	0.002245285	0.003944932	0.032885328	0.012097959	0.024740779	0.004972531	0.1713
7-T	0.054577705	0.013537056	0.039280851	0.008380739	0.251697762	0.030686941	0.184716036	0.040520680	0.0143
C-23	0.192516659	0.006676189	0.196255289	0.036512753	0.15796087	0.035997969	0.182258749	0.012709942	0.4549
C-24	0.024031565	0.005651714	0.019515219	0.004881151	0.005172879	0.006447559	0.015417974	0.001795098	0.0122
7,11-PD	0.074825949	0.003071833	0.064482453	0.011934013	0.132289453	0.033172327	0.112328337	0.006957179	0.6035
25-Br	0.000452354	5.43984E-05	0.060045876	0.014089546	0.050410424	0.040234754	0.044253929	0.006504277	0.0100
7-P	0.000452354	5.43984E-05	0.097648887	0.036678009	0.000181404	1.80986E-05	0.116786118	0.007228025	0.3195
C-25	0.090478572	0.013463302	0.092388858	0.014714305	0.054037337	0.014680537	0.076629598	0.008793739	0.1414
7,11-HD	0.538858549	0.014326582	0.407846545	0.092506721	0.309321812	0.04144878	0.229369687	0.069136724	0.3392

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A5: Average total peak area of compounds for disruption of *CG5278*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-22	1.301657918	1.789008235	1.816007309	0.124242987	2.241398297	0.053338031	3.271603223	0.687961433	0.5973
23-Br	0.043715402	0.010008862	0.042436153	0.001848504	26.67424173	11.68275607	26.43183444	12.10257215	0.1539
7-T	2.053029112	2.831589839	5.169397074	0.634716124	32.39746224	8.163567332	54.88208036	8.888468941	0.2450
C-23	25.04752605	4.487891327	22.21395429	3.634763267	58.09997649	9.910340503	51.77356967	2.663894215	0.6848
C-24	2.617598591	1.379620197	1.667490427	0.547096442	1.000085762	1.347221273	3.836610071	5.305499845	0.3245
7,11-PD	10.07098892	0.203481212	8.973224922	1.889247412	80.49983946	51.45986389	64.629002	14.93724438	0.4141
25-Br	14.98601788	0.856500855	9.513180013	3.627888879	22.66932379	1.331893794	21.38600829	1.380597549	0.6080
7-P	0.043715402	0.010008862	9.859357123	2.794803107	24.6149073	6.793396508	16.76588962	6.364521462	0.1539
C-25	11.29069416	0.360842439	9.548416599	0.643564499	20.85046026	1.664256023	15.47899941	2.083226898	0.3395
7,11-HD	73.36752211	8.043750996	61.74886336	3.318093409	102.9809472	45.77015337	94.08654332	40.41012822	0.2309

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A6: Average relative abundance of compounds for disruption of *CG5278*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-22	0.009422687	0.012964816	0.013924948	0.001274271	0.006213512	0.001461372	0.011399636	0.003727498	0.9449
23-Br	0.000311532	7.96975E-05	0.000325302	2.16954E-05	0.069978105	0.013329833	0.056809051	0.020848116	0.2393
7-T	0.014306041	0.019711526	0.039550658	0.003945506	0.087156362	0.000566193	0.183565874	0.054876214	0.0291
C-23	0.178375381	0.036806545	0.169876731	0.023905895	0.165117446	0.06928297	0.172081902	0.019884389	0.8053
C-24	0.018730697	0.010315338	0.012824567	0.004487736	0.003264704	0.004464448	0.014665583	0.017353123	0.2609
7,11-PD	0.071522901	0.000534794	0.068583562	0.012882314	0.205367929	0.085282642	0.180432192	0.008686525	0.7371
25-Br	0.106374381	0.003137707	0.07321025	0.029484833	0.063514934	0.01998387	0.077364751	0.007404958	0.1623
7-P	0.000311532	7.96975E-05	0.07578866	0.023163327	0.07088615	0.036567937	0.062216987	0.009437455	0.0900
C-25	0.080243	0.004783506	0.073101132	0.003236019	0.058576558	0.01960183	0.05151533	0.002231672	0.9961
7,11-HD	0.520401849	0.042715161	0.472814191	0.014462129	0.269924301	0.053316144	0.1899487	0.066505763	0.5862

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A7: Average total peak area of compounds for disruption of *CG6921/CG44062*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	$p^{1,2}$
C-22	1.714114627	0.822063448	1.914951481	0.659613159	6.752603547	2.769707882	3.891153033	0.581299319	0.2594
23-Br	0.051042388	0.004559301	0.051796824	0.003101615	23.61584784	3.015034374	12.80134039	2.018373366	0.0829
7-T	5.482159885	2.660337953	5.881483707	2.004463517	50.96817043	11.5665232	58.67263348	10.13743212	0.3825
C-23	21.05444666	3.863795413	28.09415569	1.542898389	74.25495597	12.45468053	53.47067934	1.636044648	0.0517
C-24	2.136746362	0.731322693	2.863113189	0.447043584	4.837169464	1.290624262	4.422197305	1.137978014	0.3814
7,11-PD	12.24660989	0.594557374	8.80832961	3.468778412	68.37069078	2.57893318	47.20679678	5.454457227	0.0654
25-Br	0.051042388	0.004559301	0.051796824	0.003101615	14.32615283	1.649125536	13.22871219	3.808978717	0.6584
7-P	20.02830011	0.471374695	28.17033963	0.665869201	15.24206711	5.074236564	27.84019992	5.854694296	0.8102
C-25	11.00159633	0.966297401	16.57089771	2.066159479	23.40053787	3.151371725	21.38403683	3.514509954	0.0077
7,11-HD	48.54487746	8.737886843	55.89977009	5.509375647	122.8812863	6.712813908	86.32153143	10.57191232	0.0301

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A8: Average relative abundance of compounds for disruption of *CG6921/CG44062*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-22	0.013843917	0.005571795	0.012742632	0.003611326	0.014383876	0.006137406	0.011788667	0.000953198	0.7831
23-Br	0.000418483	1.84174E-05	0.000349621	8.89405E-06	0.051092582	0.013737231	0.038784036	0.003814136	0.1753
7-T	0.043731233	0.018535646	0.039105933	0.010933124	0.135948524	0.006489759	0.177657037	0.020640949	0.0301
C-23	0.171387347	0.014022703	0.189660948	0.004503902	0.176834455	0.018005384	0.163228535	0.013244775	0.0730
C-24	0.017487801	0.005327754	0.019332331	0.002776369	0.011575449	0.002296829	0.013331692	0.002334897	0.9827
7,11-PD	0.100777189	0.00870879	0.058656659	0.020345045	0.158886993	0.019296672	0.143381448	0.009984979	0.1754
25-Br	0.000418483	1.84174E-05	0.000349621	8.89405E-06	0.034489742	0.001654688	0.039729615	0.007511625	0.2661
7-P	0.165694541	0.024836719	0.190498444	0.010769134	0.035937636	0.009675679	0.083997184	0.009523635	0.2201
C-25	0.091257095	0.017916426	0.112819722	0.022558399	0.063322669	0.007722959	0.065861766	0.016351426	0.3610
7,11-HD	0.394983911	0.030941591	0.376484088	0.008493815	0.317528089	0.034166896	0.26224003	0.020125359	0.2469

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A9: Average total peak area of compounds for disruption of *CG33110*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	$p^{1,2}$
C-21	4049.069444	1539.818263	678.113264	855.1768509	1289.61617	1121.963688	2036.989188	1930.26888	0.1074
C-22	216.4653207	24.72991958	316.3636775	20.49029403	739.0089039	57.11383956	823.1321063	259.6629981	0.9356
23-Br	37.0220592	1.140672245	300.0793806	59.81924784	1886.211313	135.1039588	1382.167059	334.4642128	0.0615
7-T	37.0220592	1.140672245	545.4957926	30.38666388	1580.045895	352.344589	1863.876603	17.96157577	0.5763
C-23	295.932103	319.0053405	2919.6401	2404.327568	6069.463776	1743.500162	6217.341966	4978.392744	0.5952
C-24	344.9326709	3.301071572	427.1793065	102.0271697	810.9900206	46.31762503	663.0708496	96.84004851	0.2128
7,11-PD	2029.945347	368.1505382	1999.452616	21.82465857	11513.63086	270.1479663	7959.80594	1260.574633	0.0424
7-P	1624.570735	355.5208989	1127.754705	182.5861219	2073.464173	411.1787814	1705.668641	1531.13812	0.8105
C-25	2694.19774	703.0204742	2851.414395	816.7700721	3795.159006	74.82415444	4524.911226	462.3729213	0.3457
7,11-HD	7871.014488	6619.293477	14249.71277	5489.334559	10655.81391	5529.225999	14989.56725	7913.081878	0.9143
27-Br	913.6681507	334.5117989	1547.05576	679.7904388	599.8312569	19.79351223	1265.896569	264.8042543	0.7904
C-27	4863.949198	1086.085145	3764.546471	1044.872569	1216.427484	420.7786859	2738.347168	122.8817463	0.0707

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A10: Average relative abundance of compounds for disruption of *CG33110*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-21	0.174826637	0.104292363	0.018200488	0.021184112	0.028177399	0.019873942	0.051070117	0.054770832	0.1045
C-22	0.008807586	0.001158328	0.010901348	0.003314732	0.017842634	0.002886477	0.017683927	0.001135069	0.5338
23-Br	0.001533387	0.000419701	0.010844918	0.005907785	0.045571588	0.007627328	0.029982517	0.000366715	0.0218
7-T	0.001533387	0.000419701	0.018828335	0.005887894	0.037482214	0.000561244	0.041662061	0.010186502	0.1906
C-23	0.010605668	0.010184556	0.086499078	0.046655786	0.142852255	0.007348253	0.125001666	0.076095118	0.2153
C-24	0.014249622	0.003608022	0.014246338	0.001882842	0.019627899	0.003566264	0.014564286	0.001599577	0.2734
7,11-PD	0.081909191	0.005240712	0.069583384	0.02470421	0.279778835	0.060070896	0.174557819	0.017007442	0.1228
7-P	0.065245717	0.001681683	0.040488344	0.020730122	0.049335388	0.001984047	0.042521539	0.043956193	0.6296
C-25	0.107633075	0.001891257	0.094229653	0.007834566	0.092695909	0.023793921	0.099953006	0.015357734	0.3772
7,11-HD	0.291470672	0.193909736	0.461941442	0.009931153	0.243653087	0.073046787	0.312985992	0.091938664	0.5626
27-Br	0.036017366	0.004606777	0.049618014	0.004001312	0.014559219	0.002990172	0.029082602	0.013117677	0.9338
C-27	0.206167691	0.093771815	0.124618662	0.011510218	0.028423581	0.003211368	0.060934467	0.012806186	0.1661

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A11: Average total peak area of compounds for disruption of *CG42857*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-21	256.6382658	114.210066	2140.998678	1389.830647	1347.348148	1177.035849	972.5546345	629.1557265	0.1293
C-22	37.74828899	4.757682011	876.936745	358.2585471	904.2705013	355.3816279	1288.360853	1017.530743	0.1419
23-Br	37.74828899	4.757682011	40.32835351	8.926971342	600.0827587	93.06916443	566.8011743	166.720469	0.0067
7-T	37.74828899	4.757682011	40.32835351	8.926971342	7813.453785	5415.350807	7226.209174	2933.341055	0.0170
C-23	2957.535911	950.26083	4402.264546	2969.666194	8702.29577	3732.517718	7342.361546	5826.331308	0.0985
C-24	37.74828899	4.757682011	1032.787983	342.0393205	907.8829904	304.3467686	1939.61737	1912.124179	0.7773
7,11-PD	418.8522619	426.5175384	161.1530425	44.3668716	10681.78356	3764.252356	4835.684856	1853.224571	0.0109
7-P	486.6574857	147.7299142	153.4767306	147.6928912	9425.767534	8298.753668	7744.187222	5830.116069	0.0228
C-25	2567.59757	591.3207111	3454.692984	701.3387859	3834.199198	447.6817244	4369.595343	1971.049325	0.1406
7,11-HD	11098.44346	4734.75341	13125.91188	3031.357245	10214.48163	2255.577758	10121.16109	5043.681332	0.1549
27-Br	945.5642687	175.6976803	1212.454288	269.14142	537.6468485	134.3395606	1177.658007	559.1138315	0.7465
C-27	4701.308939	896.7734116	4870.817396	2049.123212	1645.603794	327.7289286	2313.981801	489.0295913	0.6359

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A12: Average relative abundance of compounds for disruption of *CG42857*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	$p^{1,2}$
C-21	0.011282911	0.005950768	0.064101048	0.037460734	0.051278875	0.058989901	0.023402033	0.022311795	0.0260
C-22	0.001682439	0.000410652	0.027741536	0.009041118	0.014151022	0.004835766	0.024541282	0.011797615	0.0395
23-Br	0.001682439	0.000410652	0.001290989	0.000241449	0.009832589	0.004628047	0.012469257	0.004916115	0.3320
7-T	0.001682439	0.000410652	0.001290989	0.000241449	0.124029697	0.079820907	0.144397323	0.022681791	0.5836
C-23	0.125205118	0.018275646	0.135209262	0.070461124	0.154774163	0.043924695	0.130378422	0.075295682	0.5078
C-24	0.001682439	0.000410652	0.032751136	0.008032939	0.013273211	0.006965727	0.035139082	0.022697667	0.4237
7,11-PD	0.015044063	0.013180554	0.005146644	0.001233716	0.175483858	0.046507991	0.111397496	0.059364622	0.1330
7-P	0.021383313	0.007971426	0.004742779	0.003992295	0.129954259	0.099297963	0.145457634	0.056847816	0.5399
C-25	0.110675204	0.016332204	0.110269675	0.016311519	0.067770581	0.011709281	0.087798464	0.019902367	0.1807
7,11-HD	0.457376511	0.090436255	0.418691107	0.082211584	0.216155004	0.076141188	0.211718339	0.096030135	0.6640
27-Br	0.041828034	0.010998963	0.038520822	0.005865578	0.010933616	0.004514461	0.023007478	0.005749094	0.03008
C-27	0.210475093	0.058610944	0.160244014	0.074839816	0.032363125	0.009369465	0.050293191	0.016179645	0.1352

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A13: Average total peak area of compounds for disruption of *CG7485*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-22	2.370203116	1.276860569	1.693137847	0.362024602	1.030532123	1.390278931	3.083701194	0.378741	0.2586
23-Br	0.043715402	0.010008862	0.042436153	0.001848504	26.67424173	11.68275607	26.43183444	12.10257215	0.1554
7-T	2.821045717	3.917729337	5.271102073	1.462952645	32.39746224	8.163567326	54.88208036	8.888468941	0.1136
C-23	24.82738584	4.799216613	22.21395429	3.634763267	58.09997649	9.910340508	51.77356967	2.663894215	0.7930
C-24	2.226891258	1.025545355	1.876216064	0.230346818	1.59866964	2.193746712	2.746944255	1.243508392	0.8488
7,11-PD	10.07098892	0.203481209	8.973224922	1.889247412	80.49983946	15.45986388	64.629002	14.93724438	0.4637
25-Br	14.98601788	0.856500851	9.513180013	3.62788887	23.57308652	2.610007304	33.78554614	10.00400272	0.6765
7-P	0.043715402	0.010008862	9.253784759	1.938394457	25.95185061	8.684119871	21.3056518	12.7847147	0.2276
C-25	11.29069416	0.360842443	9.548416599	0.643564499	20.85046026	1.664256029	15.47899941	2.083226898	0.42392
7,11-HD	73.36752212	8.043750985	61.74886336	3.318093409	102.9809472	25.77015336	94.08654332	20.41012822	0.2342

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A14: Average relative abundance of compounds for disruption of *CG7485*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	$p^{1,2}$
C-22	0.016873994	0.009646905	0.013077435	0.003309393	0.003306404	0.004522642	0.008483212	0.000935093	0.3073
23-Br	0.000309359	8.25241E-05	0.000326647	2.73795E-05	0.069643742	0.014375187	0.069853052	0.016699495	0.2988
7-T	0.019336753	0.026826274	0.040311165	0.009615964	0.08660136	0.008441995	0.156027796	0.060257692	0.0311
C-23	0.17557364	0.040632097	0.170275428	0.021063004	0.163515496	0.066180207	0.143621687	0.026010992	0.7939
C-24	0.01582981	0.007836963	0.014393546	0.001189523	0.005141613	0.007118019	0.008067415	0.005244689	0.8224
7,11-PD	0.070924393	0.001333122	0.068716682	0.011746068	0.204759223	0.088058947	0.175532541	6.47499E-05	0.7708
25-Br	0.105462161	0.001917294	0.07372496	0.030851642	0.065871877	0.022961036	0.091052538	0.00609272	0.2112
7-P	0.000309359	8.25241E-05	0.071468015	0.017777913	0.074461964	0.041300251	0.055327324	0.021914561	0.1064
C-25	0.079595002	0.005643983	0.073333466	0.001987569	0.058048421	0.018532534	0.04252366	0.004185651	0.9756
7,11-HD	0.515785528	0.03651455	0.474372656	0.006364199	0.268649901	0.057336356	0.249510776	0.051992091	0.7321

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A15: Average total peak area of compounds for disruption of *CG8756*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-21	1552.635376	2653.645708	342.5563321	523.20695	4652.85854	6348.358378	1966.028466	2307.672199	0.6540
C-22	48.15744434	7.375028066	44.00843643	5.462837724	1445.034878	117.8853637	896.3801436	473.3671648	0.07888
23-Br	48.15744434	7.375028066	44.00843643	5.462837724	1369.043337	1181.006274	686.7106525	392.1120232	0.0034
7-T	48.15744434	7.375028066	44.00843643	5.462837724	6838.458283	2956.273568	4325.617272	1369.598181	0.0103
C-23	3528.309331	2533.738298	3972.382835	1810.978175	10280.81169	5378.500735	7730.277075	5717.866566	0.4913
C-24	48.15744434	7.375028066	44.00843643	5.462837724	1171.514954	276.9339563	645.2884475	468.8745671	0.1123
7,11-PD	1126.894495	571.5860292	913.2214434	78.67660912	12032.53303	4395.795171	4501.442108	2831.219014	0.0223
7-P	555.1959557	418.6808227	892.3248517	97.14438807	2943.074403	1362.641375	3681.439964	2494.646494	0.7505
C-25	2429.625979	1258.693131	2075.228373	719.7328051	4876.945467	1943.724626	4237.185075	1813.383667	0.8894
7,11-HD	13232.21577	8992.223129	11357.49521	3396.828131	24684.35222	16992.43538	13157.22755	5951.900984	0.3604
27-Br	1303.846737	956.4262661	1124.633828	462.3105023	1030.63838	531.5805621	1698.649623	581.3527949	0.1742
C-27	4055.348764	1661.418953	4108.940425	2906.882175	3557.044317	1475.403722	6474.588438	3995.994277	0.2341

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A16: Average relative abundance of compounds for disruption of *CG8756*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	$p^{1,2}$
C-21	0.035159011	0.049572473	0.012078705	0.017815534	0.062499297	0.078995087	0.040234515	0.047400227	0.9879
C-22	0.002381429	0.001507626	0.001958524	0.000757316	0.021036964	0.006793652	0.017840507	0.008750469	0.6295
23-Br	0.002381429	0.001507626	0.001958524	0.000757316	0.01992537	0.01825505	0.013412802	0.007712081	0.5517
7-T	0.002381429	0.001507626	0.001958524	0.000757316	0.105659511	0.061263657	0.086911492	0.019941316	0.5800
C-23	0.12078793	0.058585724	0.15606075	0.032820902	0.131743071	0.020647589	0.142685172	0.079077456	0.6537
C-24	0.002381429	0.001507626	0.001958524	0.000757316	0.016266595	0.003197227	0.013531179	0.009044205	0.6434
7,11-PD	0.050339474	0.031817452	0.039564384	0.012107374	0.163158737	0.032881929	0.087292264	0.056577436	0.1030
7-P	0.021635431	0.009551235	0.039073115	0.013488603	0.038253847	0.006982066	0.084788783	0.074012411	0.4595
C-25	0.092959414	0.018044788	0.085402196	0.023892168	0.064648165	0.005666679	0.081776407	0.023899427	0.2264
7,11-HD	0.461343406	0.151852946	0.465805369	0.100526669	0.315409124	0.116374395	0.251973622	0.078931443	0.5660
27-Br	0.043366915	0.008823495	0.045116295	0.011757085	0.013778871	0.005912547	0.034754602	0.012074404	0.0775
C-27	0.164882703	0.072483561	0.1490651	0.066449393	0.047620456	0.014212635	0.144798656	0.125140845	0.1826

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A17: Average total peak area of compounds for disruption of *CG8522*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	$p^{1,2}$
C-22	2.665270731	0.760810307	3.124156696	2.884997845	1.580883885	2.629290235	8.701000214	4.090414796	0.1137
23-Br	1.97462596	0.805873781	1.060807461	1.762532774	13.61240157	3.222512479	14.9104273	6.180903212	0.8865
7-T	4.703977485	1.848457718	6.38886873	2.010048656	105.8345879	6.851861263	130.2146557	62.6484887	0.0636
C-23	22.39440964	3.875855795	32.88589103	7.320591401	59.04502736	17.10476491	107.51236	35.56950885	0.2579
C-24	2.4249495	0.353900984	2.41405641	0.62446258	7.524285038	4.459044473	8.650179976	3.53493308	0.0189
7,11-PD	8.162480486	2.417542508	11.99606914	3.066412846	52.24317834	22.74280797	68.07649225	25.54217581	0.2386
25-Br	2.938402138	4.991552129	11.05672604	2.195605946	0.064712489	0.003942642	28.06826679	13.76428815	0.0430
7-P	0.053353135	0.006180583	16.39711998	8.193953055	0.064712489	0.003942642	72.83328345	31.29912222	0.0274
C-25	10.38395014	0.726562728	16.06670903	2.302263113	19.15205966	2.995552503	41.26991616	11.29016296	0.0988
7,11-HD	61.11192781	7.235057923	78.41559803	11.02212427	154.6230672	7.775740809	132.9754912	51.39669176	0.2571

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A18: Average relative abundance of compounds for disruption of *CG8522*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-22	0.022531214	0.00417736	0.016462543	0.014984304	0.003702907	0.00614518	0.013783895	0.001670077	0.1346
23-Br	0.016581666	0.005608015	0.005510462	0.009110481	0.033110425	0.008925236	0.024441996	0.003250939	0.7785
7-T	0.039448872	0.011418025	0.035111748	0.008258157	0.256557875	0.026714532	0.206308932	0.030261123	0.0999
C-23	0.190847531	0.013700318	0.181360054	0.025175727	0.142008567	0.03703748	0.17958429	0.016279945	0.1387
C-24	0.020842779	0.003119184	0.013299125	0.002345398	0.018511187	0.011586467	0.01404965	0.00085775	0.6746
7,11-PD	0.069049155	0.015059799	0.066176901	0.011764537	0.125100047	0.05009246	0.111819307	0.003591774	0.7458
25-Br	0.027855955	0.047443161	0.061210905	0.007719822	0.000156665	1.22368E-05	0.044492945	0.007861633	0.7069
7-P	0.00045855	5.41909E-05	0.089658431	0.038326983	0.000156665	1.22368E-05	0.117200486	0.00813091	0.2533
C-25	0.089501888	0.010699601	0.09038644	0.019028355	0.046216335	0.006052445	0.070583611	0.012892903	0.1572
7,11-HD	0.522882392	0.016100818	0.440823389	0.08946733	0.374479328	0.029228478	0.217734888	0.011235894	0.2153

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A19: Average total peak area of compounds for disruption of *CG3971*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	*p*[1,2]
C-21	2450.371808	2168.71693	635.6138434	245.0114117	4285.514214	2528.613321	1080.006908	482.9606065	0.9501
C-22	320.0671677	100.9777144	350.0280711	30.15087139	951.3789178	148.5521018	744.4357954	201.7964092	0.2442
23-Br	42.76609513	6.543947711	39.71564701	4.679610887	1409.692153	246.5461712	696.054981	56.29274399	0.0338
7-T	42.76609513	6.543947711	39.71564701	4.679610887	3210.050627	3120.287627	3038.67316	478.1339588	0.0708
C-23	2153.488708	435.5651597	1092.689876	994.0319239	8231.960558	958.8904359	5436.595065	2799.397524	0.1763
C-24	458.2309369	39.63331706	556.2520185	44.33773482	1115.011143	214.9153965	858.9126164	326.6913086	0.24081
7,11-PD	2419.245104	1458.042675	1565.92139	216.8595926	7222.355992	1741.509458	5671.191461	563.7937602	0.3757
7-P	2934.352417	2351.559941	1365.62405	30.39155292	1688.135286	29.03675586	3952.745546	831.4443876	0.3029
C-25	2977.501347	258.3601681	2636.894146	275.5115095	3395.190254	1348.027847	4701.40217	1144.042642	0.3749
7,11-HD	16456.40208	13964.5155	6438.823537	1254.30147	7122.919516	3872.429285	11562.22058	1883.509968	0.9158
27-Br	1018.982651	60.28611829	948.9085828	100.7006297	1082.773774	108.1961254	1041.283732	203.2997934	0.3681
C-27	5898.018589	490.0690867	6324.83313	712.1986201	1791.86231	104.8195735	3023.160073	814.8873887	0.5195

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Table A20: Average relative abundance of compounds for disruption of *CG3971*.

Compound	Mel/Bal[1]	SD	Mel/Dis[1]	SD	Sim/Bal[1]	SD	Sim/Dis[1]	SD	$p^{1,2}$
C-21	0.058765761	0.025981796	0.029896452	0.015094256	0.055378609	0.028019372	0.027667771	0.017389552	0.9724
C-22	0.009267598	0.002386921	0.015962971	0.000740857	0.023485487	0.002820941	0.017691172	0.001094429	0.0108
23-Br	0.001298963	0.000539259	0.0018358	0.000455607	0.028029649	0.012957851	0.016883081	0.002215473	0.2777
7-T	0.001298963	0.000539259	0.0018358	0.000455607	0.047959636	0.025091787	0.073104468	0.003986722	0.2427
C-23	0.072089183	0.051414991	0.047101882	0.038962628	0.173345761	0.076475053	0.125776963	0.040424239	0.7819
C-24	0.014184307	0.006744691	0.025378631	0.001341404	0.027806793	0.001678697	0.020169427	0.003559008	0.0282
7,11-PD	0.063983503	0.003990095	0.072480673	0.019447591	0.139732668	0.057463383	0.137285584	0.015478633	0.8172
7-P	0.072516245	0.023328602	0.062728117	0.009679735	0.041542023	0.012252164	0.094554632	6.12008E-05	0.0340
C-25	0.0921598	0.043799456	0.120108364	0.003362448	0.127884409	0.004285309	0.112063116	0.003722112	0.2347
7,11-HD	0.399905111	0.155454487	0.291518872	0.018463041	0.233470691	0.143490333	0.277998486	0.013598983	0.3669
27-Br	0.031785225	0.015881425	0.043217288	0.001138661	0.035517807	0.007385962	0.024949408	0.000400945	0.1510
C-27	0.182745341	0.087448979	0.287935149	0.005709504	0.065846467	0.019984682	0.071855893	0.004331718	0.1940

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the interaction term.

Appendix B: Average total peak area and relative abundance of compounds in microbial and diet experiments

Table B21a: Average total peak area for compounds across different microbial treatments

Compound	WT[1]	Ab[1]	Re[1]	Act[1]	Lac[1]	AL[1]	*p*[1,2]
7-T	2.638559578	3.500658452	2.322500122	2.019668601	2.066895877	2.224742661	0.0215
C-23	13.33369949	7.720498279	7.194778755	6.58732764	7.484761049	6.142216675	0.0966
6.82	14.24256869	15.22927868	14.45087358	14.67446919	15.46389912	15.24225829	0.9588
C-24	2.303596441	2.343949323	2.654187161	2.976672836	2.559918859	3.095863785	0.1136
7,11-PD	8.690257431	8.894918376	8.341649004	9.577814296	8.355588103	9.897478037	0.8210
25-Br	11.98105437	11.55258061	9.631078833	10.54025468	9.485488605	10.39617116	0.7627
7-P	8.359083029	10.3020166	9.035678807	10.58396619	9.205302388	10.74601888	0.4608
C-25	10.36899813	11.6019637	10.70138126	13.84803177	11.74463516	11.4310775	0.3116
7,11-HD	65.01286746	72.62943201	65.25117784	73.25410428	63.60148092	73.06737347	0.7444
27-Br	9.724613699	10.07499553	9.096681777	12.53054535	8.426737861	9.40955485	0.0518
10.09	11.81721053	13.03416974	13.78889116	14.54041544	14.06637608	14.20924402	0.4771
9-H	9.433610948	9.774222175	9.344265071	12.6661994	10.04231616	11.90108908	0.3489
C-27	2.781075135	2.62888924	2.734496877	2.708974192	2.449764665	2.474893421	0.9801
29-Br	53.9467929	55.52403894	50.70108928	65.95309343	54.80459985	56.68815664	0.1494

[1]Bolded entries pass the FDR cutoff.

[2] The *p*-value is for the treatment (microbial type) term.

Table B21b: SD Values for total peak areas from table B21a

Compound	SD WT	SD Ab	SD Re	SD Act	SD Lac	SD AL
7-T	0.95112989	1.659450948	0.912616252	0.730606529	0.394771812	0.409106745
C-23	12.64196736	3.515998187	2.260585368	2.323972102	3.90811769	1.774694514
6.82	2.29337838	4.67759229	3.004763239	4.963780467	2.445862934	2.810711622
C-24	0.418644519	1.02427118	0.277992728	0.912042684	0.64028526	0.426639513
7,11-PD	3.050322562	3.741274391	3.511824144	2.447533564	2.450782384	2.786266701
25-Br	6.927663599	5.386333004	2.541143227	3.794018641	2.164434378	1.787478621
7-P	3.703986936	5.054153839	2.24692985	2.451921781	2.234178443	2.413628092
C-25	3.083382737	5.158989417	2.682411644	3.452808757	2.695798316	1.431746737
7,11-HD	41.50148031	31.44510262	25.18600088	17.9284164	11.0972504	9.412791414
27-Br	5.136716658	3.79332619	2.423558013	3.985986101	1.097200704	3.462410144
10.09	3.154442072	2.099105088	4.249747788	3.847840168	2.744741435	2.087064648
9-H	2.77743117	5.013872606	2.948616384	2.540808488	2.153557232	5.097707949
C-27	2.009589901	1.859857056	1.00796877	1.565564963	0.377682168	0.567370589
29-Br	23.68084054	19.13770996	14.73829481	18.32939056	6.378590771	11.85721723

Table B22a: Average relative abundance of compounds across different microbial treatments

Compound	WT[1]	Ab[1]	Re[1]	Act[1]	Lac[1]	AL[1]	*p*[1,2]
7-T	0.0108027	0.014774576	0.010636895	0.007966497	0.009449925	0.009596153	0.0414
C-23	0.043158843	0.032489148	0.03357913	0.026329848	0.03306367	0.02651799	0.2109
6.82	0.066718758	0.065142327	0.068572338	0.059087855	0.070315587	0.064498364	0.5304
C-24	0.011820417	0.010127855	0.013069706	0.011698328	0.011779029	0.013317538	0.4485
7,11-PD	0.040161682	0.043109665	0.037591043	0.038488911	0.037537015	0.041348668	0.8378
25-Br	0.056848696	0.046700808	0.045309391	0.041153197	0.043206967	0.044403058	0.1792
7-P	0.035162945	0.042974547	0.042422729	0.042853316	0.041630891	0.045378458	0.34309
C-25	0.054157731	0.050594524	0.050310934	0.054722688	0.053219446	0.049013697	0.9211
7,11-HD	0.27889936	0.305828804	0.295795699	0.28936395	0.288774234	0.309188076	0.7935
27-Br	0.045950919	0.042082409	0.043073814	0.049035206	0.038767986	0.038960055	0.0153
10.09	0.059659277	0.059476001	0.066057518	0.05796437	0.063881356	0.060448507	0.8456
9-H	0.044395328	0.04346292	0.045342144	0.050572	0.045773502	0.048832622	0.8984
C-27	0.012892573	0.009828944	0.012851562	0.010356587	0.011397364	0.010457345	0.3865
29-Br	0.239370771	0.233407472	0.235387097	0.260407247	0.25120303	0.238039469	0.3145

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the treatment (microbial type) term.

Table B22b: SD Values for relative abundances from table 22a

Compound	SD WT	SD Ab	SD Re	SD Act	SD Lac	SD AL
7-T	0.004417096	0.007545484	0.002316952	0.00230601	0.001629342	0.002315567
C-23	0.025031068	0.010433607	0.006020691	0.008488361	0.012402528	0.008295666
6.82	0.019285917	0.008083313	0.008894713	0.01857899	0.006129645	0.008042672
C-24	0.004512574	0.004651147	0.003148342	0.001896352	0.003034699	0.00277495
7,11-PD	0.01507428	0.009913167	0.007078922	0.009884192	0.007755251	0.007561577
25-Br	0.028981226	0.00794601	0.005976905	0.009135325	0.008595328	0.007984555
7-P	0.010915085	0.012198998	0.003746193	0.009457388	0.007924927	0.007995806
C-25	0.027812761	0.014378295	0.005985873	0.006227649	0.007371161	0.008783452
7,11-HD	0.117246804	0.040829171	0.034719336	0.025481642	0.022673538	0.016872275
27-Br	0.008061652	0.006088058	0.008151136	0.004973207	0.00579613	0.010544573
10.09	0.023393781	0.018353588	0.018095554	0.012573939	0.006694053	0.00808487
9-H	0.01573742	0.019271201	0.016646416	0.006379943	0.007559203	0.015735993
C-27	0.005147269	0.004900635	0.003337278	0.004540701	0.002689401	0.001893499
29-Br	0.049176552	0.025375068	0.013928812	0.027184434	0.025404929	0.025939824

Table B23: Average total peak areas for compounds across different diet types

Compound	CMY[1]	SD[1]	HY[1]	SD[1]	*p*[1,2]
7-T	2.422986604	0.188318073	2.50135516	0.233469388	0.7643
C-23	9.624746719	1.598082422	6.529680577	0.541625297	0.0477
6.82	13.85082882	0.680384279	15.91695369	0.639691212	0.0270
C-24	2.526107284	0.156270689	2.785288851	0.127752476	0.1830
7,11-PD	8.129343245	0.563745339	9.789891837	0.595764228	0.0449
25-Br	10.19949808	0.907466185	10.99604468	0.759363812	0.4899
7-P	9.036115341	0.614993699	10.37457329	0.653161262	0.1088
C-25	11.08645034	0.617366659	12.14557884	0.721615617	0.24395
7,11-HD	**55.94202037**	**4.520089821**	**81.66345829**	**3.98980168**	**1.44301E-05**
27-Br	**7.834096363**	**0.539812101**	**11.92027999**	**0.671851731**	**2.92575E-06**
10.09	12.86994048	0.627857902	14.28216184	0.62013836	0.1053
9-H	10.23657905	0.671569449	10.8173219	0.825622804	0.5892
C-27	**1.8887464**	**0.195255307**	**3.370618109**	**0.245906854**	**1.59179E-05**
29-Br	**46.27052608**	**2.342245485**	**66.26873094**	**2.969476886**	**3.01822E-07**

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the diet type term.

Table B24: Average relative abundance of compounds across different diet types

Compound	CMY[1]	SD[1]	HY[1]	SD[1]	$p^{1,2}$
7-T	0.011652813	0.001052462	0.009422769	0.000614756	0.0646
C-23	0.040132434	0.002980484	0.024913775	0.0015306	0.2103
6.82	0.069709674	0.003015908	0.061735402	0.001687276	0.0276
C-24	0.013028003	0.000827629	0.010909621	0.000480334	0.0303
7,11-PD	0.042043	0.002266133	0.037369328	0.001498069	0.1104
25-Br	0.050831511	0.003465582	0.041709195	0.001758671	0.0141
7-P	0.043551633	0.001927327	0.039922662	0.001783881	0.1729
C-25	0.056865822	0.003001613	0.047140519	0.002091551	0.0105
7,11-HD	**0.274161785**	**0.013298344**	**0.315121589**	**0.005019066**	**0.0044**
27-Br	**0.040345802**	**0.001693681**	**0.045610994**	**0.001406269**	**0.0060**
10.09	0.066377072	0.003229856	0.056118605	0.002594201	0.01386
9-H	0.050655662	0.002615999	0.042137177	0.002804883	0.0382
C-27	**0.009756639**	**0.000807413**	**0.012838152**	**0.000669632**	**0.0052**
29-Br	**0.230888151**	**0.006179863**	**0.255050211**	**0.004958536**	**0.0036**

[1]Bolded entries pass the FDR cutoff.
[2] The *p*-value is for the diet type term.

Curriculum Vitae

Name:	Heather Ward
Post-secondary Education and Degrees:	University of Waterloo Waterloo, Ontario, Canada 2010-2015 B.Sc. The University of Western Ontario London, Ontario, Canada 2015-2017 (in progress) M.Sc
Honours and Awards:	Natural Sciences and Engineering Research Council (NSERC) CGS-M 2016-2017 Ontario Graduate Scholarship (OGS) (Declined) 2016-2017
Related Work Experience	Teaching Assistant University of Waterloo 2014 Teaching Assistant The University of Western Ontario 2015-2017

Presentations:

Ward, H. and Moehring, A.J. Evolution 2017, Portland, OR. *The genetic basis for cuticular hydrocarbon biosynthesis in* Drosophila. Oral Presentation.

Ward, H. and Moehring, A.J. BGRF 2017, The University of Western Ontario. *The genetic and environmental basis for CHC production in* Drosophila. Lightning Talk Oral Presentation.

Ward, H. and Moehring, A.J. OE3C 2016, University of Toronto. *The genetic and environmental basis for cuticular hydrocarbon biosynthesis in* Drosophila. Lightning Talk Oral Presentation.