Gene Expression and Introgression of Two Genes Implicated in Behavioral Reproductive Isolation between *Drosophila simulans* and *D. sechellia*

By

Denny R. Swartzlander

Submitted to the graduate degree program in Molecular Biosciences and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Chairperson: Jennifer M. Gleason

_____

Erik A. Lundquist

_____

Stuart J. Macdonald

_____

Robert E. Ward

_____

Justin P. Blumenstiel

Date Defended: 9 December 2016

The dissertation committee for Denny R. Swartzlander certifies that this is the approved version of the following dissertation:

Gene Expression and Introgression of Two Genes Implicated in Behavioral Reproductive Isolation between *Drosophila simulans* and *D. sechellia*

_____

Chairperson: Jennifer M. Gleason

Date Approved: 15 December 2016

Abstract

Reproductive isolation maintains species barriers and can cause population divergence. Pre-mating isolation prevents courtship between species and can be caused by changes in gene expression resulting in sex- and species-specific phenotypes. In *Drosophila*, courtship is heavily influenced by the production and reception of cuticular hydrocarbons (CHCs), which act as pheromones that elicit or repel courtship from a potential mate. Differences in expression of genes involved in CHC biosynthesis between species can produce different types of CHCs, resulting in species-specific CHC profiles. Differential CHC profiles can result in courtship barriers between species.

Drosophila simulans and *D. sechellia* are asymmetrically reproductively isolated from each other in part because of differential production of CHCs. *D. simulans* males and females produce 7-tricosene (7-T) whereas *D. sechellia* males produce 7-T and *D. sechellia* females produce 7,11-heptacosadiene (7,11-HD). 7,11-HD acts as an anti-aphrodisiac to *D. simulans males*, which only court *D. simulans* females. *D. sechellia* males court both *D. simulans* and *D. sechellia* females. Thus reproductive isolation occurs between *D. simulans* males and *D. sechellia* females.

Genomic regions containing *desatF* and *eloF*, among other genes, have been identified by quantitative trait locus (QTL) studies as potentially contributing to production of the *D. sechellia* pheromone 7,11-HD. In this study I tested the hypothesis that *desatF* and *eloF* influence reproductive isolation between *D. simulans* and *D. sechellia*. In the first set of experiments, using gene expression analysis, I measured the mRNA production of *desatF* and *eloF*, as well as other desaturases and elongases associated with QTL affecting CHC differences between *D. simulans* and *D. sechellia*. Both *eloF* and *desatF* were expressed only in females of *D. sechellia*. The

other genes had variable expression patterns that did not suggest involvement in sex-specific CHC production. Using allele-specific qPCR in *D. simulans*/*D. sechellia* hybrids, I found that only the *D. sechellia* alleles of *desatF* and *eloF* are expressed, implying that *desatF* and *eloF* expression differences between females of *D. simulans* and *D. sechellia* are likely caused by a *cis*-regulatory change.

To further examine the differences in *desatF* and *eloF* expression, in the second set of experiments I introgressed the *D. simulans* alleles of *desatF* and *eloF* into a *D. sechellia* background. Introgression lines did not express the *D. simulans desatF* and/or *eloF* alleles. To determine the effect of *desatF* and *eloF* expression on courtship, I measured the courtship behavior of *D. simulans* and *D. sechellia* males toward the introgression lines. The time required for a male to start courting did not depend upon the target females, whether it was an introgression line, conspecific or heterospecific. In contrast, the courtship effort of males differed by target female. Male *D. simulans* courted all introgression lines, and some lines received significantly more courtship than *D. sechellia* females, though not as much as *D. simulans* females, indicating that the change in allele removed some of the reproductive barrier between *D. simulans* males and *D. sechellia* females. The type of female had less of an effect for *D. sechellia* males, which was expected because *D. sechellia* males are not completely reproductively isolated from *D. simulans* females. Altogether these results imply that *desatF* and *eloF* expression differences between *D. simulans* and *D. sechellia* directly affect reproductive isolation between the species, presumably through the genes' roles in 7,11-HD production.

Acknowledgments

I thank my P.I. Jennifer Gleason for her patience and understanding of the difficulties with being in graduate school while also having children. Having three kids born while going through graduate school was difficult on many levels, but also of course infinitely rewarding. Jenny was extremely helpful in giving me the time and support needed to finish graduate school while I raised children and cared for my family. I also learned a great deal from Jenny about how to perform scientific research and how to write in a clear and effective scientific manner. Her help and mentorship was very important in allowing me to reach the completion of my graduate career.

I thank Kaila Colyott and Paula Roy, both fellow graduate students, for their support over the years of getting through the everyday stress of graduate school.

I thank the undergraduate researchers for their much appreciated help in lab work, especially the seemingly endless genotyping. In particular, I thank Leah Logan for her valuable help in building my near-isogenic lines.

I thank my committee, which included Stuart Macdonald, Erik Lundquist, Rob Ward and Justin Blumenstiel, for their advice and continued support of me reaching my goals over the years.

I thank my parents and parents-in-law for their constant support and encouragement. I also thank my wife Morgan for her love and support, and for caring for our children while I spent long hours away in the lab or at the library writing. And finally, I thank my children Mary, Benjamin, and Peter for being constant sources of happiness and reminding me to have fun.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction to Reproductive Isolation, Gene Expression, and Courtship in *Drosophila*

Introduction

Reproductive isolation, either through premating or postmating isolation, is instrumental in maintaining species boundaries (Dobzhansky 1935). Premating isolation involves mechanisms causing the failure of different species to mate with one another. One such mechanism is the failure of courtship. Postmating isolation is caused by gametic incompatibility, hybrid sterility and hybrid inviability (Orr et al. 2004). While studies on the genetics of postmating isolation in *Drosophila* have characterized genomic incompatibilities (e.g. Ting et al. 1998; Sawamura et al. 2004; Fang et al. 2012), little is known about the genetics of traits influencing sexual isolation through behavior (reviewed in Etges 2014).

Understanding the genetics of chemosensory signaling and mating behavior traits in *Drosophila* can be informative for understanding reproductive isolation, and ultimately speciation (Smadja and Butlin 2009). Still, few studies have investigated how speciation can be affected by chemically influenced behavior (reviewed in Smadja and Butlin 2009). Among species of *Drosophila*, behavioral changes influencing population divergence can occur as a result of alterations to multiple sensory modalities, including chemical signals such as pheromones involved in courtship (reviewed in Smadja and Butlin 2009). Such alterations to pheromone biosynthesis pathways can occur because of changes in gene regulation that affect gene expression and produce species-specific phenotypes (Mackay 2004; Shirangi et al. 2009; Wittkopp and Kalay 2012). Gene expression differences in pathways producing pheromones can also lead to sex-specific phenotypes within *Drosophila* species (Chertemps et al. 2006; Chertemps et al. 2007).

Courtship in Drosophila

Courtship behaviors in *Drosophila* are well documented and the steps involved in courtship are well understood (e.g. Spieth 1974; Stocker and Gendre 1989; Tomaru et al. 1998; Tauber and Eberl 2002). In *Drosophila melanogaster*, courtship begins with a male approaching a female and orienting towards her. The male then taps her abdomen with his forelegs. Gustatory receptors located on the male's forelegs detect chemical signals present on the female's abdomen (Cobb and Jallon 1990; Ferveur 2005; reviewed in Montell 2009). Such chemical cues relay information to the male about the individual he is courting, such as the individual's sex, species, and mating status (Cobb and Jallon 1990). An experienced male who detects a female of both the correct species and mating status continues courtship by producing a species-specific acoustic signal through wing vibration (reviewed in Kyriacou and Hall 1982; Ritchie et al. 1998). Other courtship signals include chasing of the female by the male, licking of the female genitalia by the male, and female spreading of wings as a sign of acceptance to the male. Copulation occurs if both the male and female accept each other's courtship signals.

Cuticular hydrocarbons and their effect on courtship

Drosophila courtship is influenced by the presence of cuticular hydrocarbons (CHCs) secreted onto the cuticle of the fly (Ferveur 1997). Oenocytes, secretory cells found in the *Drosophila* abdominal segments, produce CHCs primarily for desiccation resistance (Rouault et al. 2004). However, CHCs also function as contact pheromones (Ferveur 2005). The progression of courtship can be halted or encouraged depending on the female CHCs detected by the male. If the male finds the female's CHCs attractive, courtship may continue. Conversely, if the male

finds the female CHCs aversive, further courtship may be prevented. Thus, the type of CHC produced by the female affects whether or not she is courted.

Female *D. simulans* and *D. sechellia* produce different CHCs (Cobb and Jallon 1990). The predominant CHC for both sexes of *D. simulans* is 7-tricosene (7-T, 23:1), and is known as the monomorphic state. 7-T has 23 carbons with a single double bond at the 7^{th} carbon. *D. sechellia* CHC production is sex-specific. The predominant CHC of *D. sechellia* males is 7-T. Females have very little 7-T and instead the predominant female CHC is 7,11-heptacosadiene (7-11, HD; 27:2; Cobb and Jallon 1990). This sex-specific condition is known as the dimorphic state. 7,11-HD has 27 carbons with double bonds at the 7^{th} and the 11^{th} carbons. Thus, the predominant female CHCs of these two species differ in both chain length and the number of double bonds.

The production of 7,11-HD by *D. sechellia* females is a major contributing factor to a courtship barrier between the two species (Cobb and Jallon 1990). 7,11-HD is an anti-aphrodisiac for *D. simulans* males (Billeter et al. 2009), which do not court *D. sechellia* females (Coyne et al. 1994). *D. sechellia* males court females of both species (Jallon 1984) and the absence of 7,11-HD in *D. simulans* females does not prevent courtship (Billeter et al. 2009). Thus, an asymmetric reproductive barrier exists between these two species.

In addition to 7-T and 7,11-HD, other CHCs differ between these two species in number of carbons and double bonds (Jallon and David 1987). The most striking CHC differences between the two species occur between the females. The general trend in CHC structure is that *D. sechellia* females have CHCs with longer, less saturated compounds than those of *D. simulans* females.

CHC profiles in the D. melanogaster group

CHC profiles of species within the *D. melanogaster* group vary by both species and sex. The sex-specific CHC profile of *D. melanogaster* is the same as it is in *D. sechellia* (7-T in males; 7,11-HD in females). *D. erecta* has a similar profile in that the predominant female CHCs are longer and less saturated than those of the males (Jallon and David 1987). The predominance of 7-T in both sexes of *D. simulans* is the same in the sister species, *D. mauritiana*. These patterns of CHC production are correlated with courtship. Males from sexually monomorphic species court females of other sexually monomorphic species (Cobb and Jallon 1990). In contrast, males from sexually dimorphic species, in which females produce dienes, preferentially court females that produce dienes, although they will court females without dienes (Billeter et al. 2009). This courtship barrier between monomorphic males and dimorphic females can be affected by artificially altering CHCs. Applying 7,11-HD to *D. simulans* females, which normally produce 7-T, causes a failure of courtship from *D. simulans* males. Removing 7,11-HD from *D. melanogaster* females can elicit courtship from *D. simulans* males, which normally would not court *D. melanogaster* females. Also, applying 7-T to *D. sechellia* females can make those females attractive to *D. simulans* males, because 7-T is the attractant normally present on *D. simulans* females. Thus, CHC profiles play a large role in mate selection.

Because the relationships among the three sibling species (*D. simulans, D. sechellia* and *D. mauritiana,* Figure 1.1) are not well established (Garrigan et al. 2012b), the production of 7,11-HD in *D. sechellia* females may represent a loss followed by an independent gain of sexual dimorphism, or by contrast, *D. simulans* and *D. mauritiana* may have independent losses of sexual dimorphism. Thus, examining the genetic basis of CHC production across species will determine whether these are homologous or convergent phenotypes with respect to *D. erecta*.

Similar CHC biosynthesis pathways among *D. erecta*, *D. melanogaster*, and *D. sechellia* suggest homology, while a unique pathway in *D. sechellia* compared to *D. erecta* and *D. melanogaster* suggests the convergence of CHC phenotypes.

CHC production pathway in D. melanogaster

The CHC biosynthesis pathways in *D. simulans* and *D. sechellia* are currently unknown, however the biosynthesis pathways of *D. melanogaster* CHCs have been characterized (Legendre et al. 2008). In *D. melanogaster*, biosynthesis of CHCs starts from a common precursor molecule that is modified by elongation (adding carbons in pairs), desaturation (removing hydrogens to form double bonds), and decarboxylation (removing single carbons, Figure 1.2, Legendre et al. 2008). Female CHCs are produced by further modification that does not occur in the production of male CHCs (Legendre et al. 2008). Two enzymes are involved in elongation and desaturation in *D. melanogaster* females: (1) DESATF, a desaturase; and (2) ELOF, an elongase. The genes for DESATF and ELOF are not expressed in *D. melanogaster* males, but are expressed in females, leading to sexual differentiation of CHCs (Chertemps et al. 2007; Legendre et al. 2008). Two other desaturases (DESAT1 and DESAT2) involved in *D. melanogaster* CHC production are not sex-specific, but the production of DESAT2 is found primarily only in African populations (Dallerac et al. 2000).

Evolution of gene expression

The expression of a gene, or lack thereof, is essential for resulting in a phenotype that can be acted on by selection. CHC type in *D. melanogaster* is one such phenotype affected by the expression of desaturase and elongase genes. The expression differences between *D. melanogaster* males and females result in CHC pheromone phenotypes that can cause

reproductive isolation from other species. One cause for the rise of such sexually dimorphic traits is a change in gene expression between the sexes. Thus, understanding gene expression is important to understanding the genetic basis of sexually-selected traits.

Gene expression is controlled by both *trans*-acting and *cis*-acting factors. *Trans*-acting factors are proteins such as transcription factors that bind to specific *cis*-regulatory sequences. *Cis*-acting factors are DNA sequences such as transcription factor binding sites and promotor elements of genes targeted for transcription.

Cis-acting changes in gene expression are more likely than *trans*-acting changes to result in novel traits causing reproductive isolation (Wittkopp et al. 2004; Landry et al. 2005; Wittkopp et al. 2008b, a; Tirosh et al. 2009; Emerson et al. 2010; McManus et al. 2010; Wittkopp and Kalay 2012; Coolon et al. 2013; Coolon et al. 2014; Meiklejohn et al. 2014). *Cis*-regulatory changes are more common among species than within species (Wittkopp et al. 2004), and expression differences attributable to *cis*-regulatory changes are greater among than within species (Wittkopp et al. 2008b; Tirosh et al. 2009; Emerson et al. 2010). *Cis*-factors have a larger effect than *trans*-factors on overall expression levels (Zhang et al. 2011; Gruber et al. 2012; Meiklejohn et al. 2014), and are hypothesized to have less pleiotropic effects than *trans*-acting changes, which can have negative effects on many different genes (Stern 2000).

Adaptive evolution of biosynthesis pathways affecting complex behavior can arise through mutations in both regulatory and protein coding sequences (reviewed in Molodtsova et al. 2014). Changes in DNA regulatory sequences may affect the amount of gene transcripts present, while changes in protein coding sequences may affect protein structure and consequently function. Adaptive *cis*-regulatory sequence changes may play a larger role than coding changes in the evolution of novel complex traits (Wray 2007; Carroll 2008; Garfield and

Wray 2010). While both types of mutations can affect biosynthesis pathways, *cis*-regulatory sequence mutations are more likely to alter the expression of a single gene in a pathway, and coding region changes to proteins such as transcription factors are more likely to change the regulation of large sets of target genes (Cheatle Jarvela and Hinman 2015). Coding region changes that alter protein function in a pathway could also be less biologically tolerated if that protein is involved in other tissue-specific pathways. In support of this idea, the coding regions of genes with pleiotropic effects are often conserved (Cheatle Jarvela and Hinman 2015). Thus, coding region changes can have more widespread effects outside a pathway, while the effects of *cis*-regulatory changes tend to be confined to the affected pathway, making changes in *cis*-regulation more likely to lead to adaptive evolution of biosynthesis pathways (Cheatle Jarvela and Hinman 2015). If *cis*-acting regulatory differences have a larger effect on reproductive isolation, they could play a more important role in speciation than *trans*-acting regulatory differences. Novel *cis*-acting regulatory variation could be more likely to cause phenotypes that introduce reproductive isolation between subsets of populations (Wittkopp et al. 2004).

While *cis*-regulatory sequence variation is considerable among species (Borneman et al. 2007; Bradley et al. 2010), changes in *trans*-acting factors, such as transcription factors, may play a larger role in evolution than previously thought (Yvert et al. 2003; Bustamante et al. 2005; Wagner and Lynch 2008; Lynch et al. 2011). Changes in *trans*-acting factors can compensate for and alter the severity of *cis*-acting changes (Coolon et al. 2014). Also, changes in transcription factor protein sequences appear to be heavily involved in the regulation of gene expression, because sequence changes in transcription factor DNA binding domains can alter binding specificity (reviewed in Weirauch et al. 2014). Thus, changes in both *cis*- and *trans*- effects are important in the evolution of gene regulation, although the ways in which they affect regulation

can be quite different (Gordon and Ruvinsky 2012). Understanding the evolution of gene regulation can reveal how closely related species have evolved species-specific courtship behaviors that can maintain reproductive isolation, and ultimately speciation.

Desaturase activity in Drosophila

Desaturase enzymes have many functions in *Drosophila* physiology, particularly in cellular functions including lipid metabolism, response to autophagy, response to sucrose (Kohler et al. 2009), glucose homeostasis (Musselman et al. 2013), regulation of cell size, regulation of starvation and imaginal disc growth (Parisi et al. 2013), and are expressed in a variety of tissues and at different developmental times (Tomancak et al. 2002; Billeter et al. 2009; Weiszmann et al. 2009; Frise et al. 2010). In *D. melanogaster* CHC production, desaturases modify long-chain fatty acids by removing hydrogen atoms (Legendre et al. 2008). At least ten desaturases are known in *Drosophila* (reviewed in Gleason et al. 2009) and of those, three (*desatF*, *desat1*, and *desat2*) are known to play a role in CHC production (Dallerac et al. 2000; Labeur et al. 2002; Legendre et al. 2008). In addition, *desat1* is known to play a role in the perception and emission of pheromones in neural and non-neural tissues (Bousquet et al. 2012).

Desaturases belong to a gene family that has undergone gene duplications, deletions, and diversifying selection (Fang et al. 2009; Keays et al. 2011). Desaturase loci involved in pheromonal function have a higher rate of duplication and loss than other loci within the same gene family involved in non-pheromonal functions (Keays et al. 2011). Additionally, in some *Drosophila* species strong positive selection has been detected on desaturase coding sequence divergence, yet desaturase activity has been retained (Keays et al. 2011). Thus, desaturases specifically involved in CHC production appear to be under strong selection pressure that conserves desaturase function.

The mRNA expression of *desatF* (syn. *Fad2*) correlates with the production of CHC dienes in the subgenus *Sophophora* (Shirangi et al. 2009). The *cis*-regulatory sequence elements (CREs) controlling *desatF* expression have evolved rapidly within the *melanogaster* subgroup. Both sexes of species outside the *melanogaster* subgroup produce dienes but, within the *melanogaster* subgroup, diene production in *D. melanogaster* and *D. sechellia* is female specific. Additionally, neither sex of *D. simulans* or *D. mauritiana* produces dienes. Gene expression of *desatF* has the same sex and species pattern as diene production. Females of species that produce dienes express *desatF*, whereas males of those species do not. In species that do not produce dienes in either sex, *desatF* is also not expressed in either sex.

Sex- and species-specific expression of *desatF* is governed by CREs. One known CRE for *desatF* is a binding site for the transcription factor DOUBLESEX (DSX). In *D. melanogaster* and *D. erecta*, the female isoform of DSX binds to a *desatF* CRE, activating transcription (Shirangi et al. 2009). However, a DSX binding site is missing in *D. sechellia*, suggesting the possibility of other transcription factors activating the *desatF* expression pathway (Shirangi et al. 2009). In addition, the *desatF* DSX CREs in *D. melanogaster* and *D. erecta* are not identical. This suggests that the DSX binding site controlling *desatF* expression has undergone evolutionary changes in the melanogaster group. Still, *desatF* is expressed in females of all three species (Shirangi et al. 2009). Furthermore, another *desatF* CRE in *D. melanogaster*, a possible transcription factor binding site, appears to have lost a small number of nucleotides in comparison to the same sites in *D. erecta* and *D. simulans*, which have nearly identical sequences for that CRE (Shirangi et al. 2009). Thus the CRE evolved in *D. melanogaster* through a series of small deletions that did not occur in *D. erecta* or *D. simulans*. These findings

suggest that *desatF* is under stabilizing selection maintaining expression in dimorphic species while allowing CRE sequence alterations.

Furthermore, in the case of *D. melanogaster desatF*, *cis*-regulatory mutations can account for alterations to CHC structure. Absence of *desatF* expression results in a lack of dienes (Legendre et al. 2008). *desatF* must be expressed for a CHC structure to contain more than one double bond (Legendre et al. 2008). Thus, in *D. melanogaster*, *cis*-regulatory mutations that affect *desatF* expression are sufficient to either produce or prevent diene formation in CHCs.

Elongase activity in Drosophila

Elongases are less studied than desaturases in *Drosophila*, but are known to be involved in pheromone metabolism, fatty acid biosynthesis and elongation (Chertemps et al. 2007; Ng et al. 2015). Within the *Drosophila* genome, elongases are often clustered in close proximity to one another and have similar DNA sequences, suggesting a history of gene duplications (Attrill et al. 2016). Of the 19 identified potential genes in the *D. melanogaster* elongase gene family, only three have been functionally characterized (Chertemps et al. 2005; Chertemps et al. 2007; Ng et al. 2015). *elo68a* is specifically expressed in the testis and ejaculatory bulb of the male reproductive system, and is involved the biosynthesis of the nonhydrocarbon male pheromone *cis*-vaccenyl acetate (Chertemps et al. 2007). *bond* is another elongase also expressed in the ejaculatory bulb, and is involved in the production of the male pheromone (3*R*,11*Z*,19*Z*)-3-acetoxy-11,19-octacosadien-1-ol (CH503) and in fertility (Ng et al. 2015). Both cis-vaccenyl acetate and (3*R*,11*Z*,19*Z*)-3-acetoxy-11,19-octacosadien-1-ol (CH503) are deposited by the male into the female during copulation to discourage copulation from other males. *eloF* is the only elongase known to be involved in CHC production, elongating female dienes in *D. melanogaster* (Chertemps et al. 2007). The activity of *eloF*, as well as the relatively sparse functional

knowledge of the elongase family, makes elongases attractive candidates for research into their roles in CHC production.

The expression of *eloF* in *D. melanogaster* is specific to females, and affects CHC length and courtship behavior (Chertemps et al. 2007). RNAi knockdown of the *eloF* mRNA transcript shortens CHC length in females, and results in decreased female attractiveness affecting wild-type male courtship (Chertemps et al. 2007). The regulation of *eloF* expression is not well understood, and no *cis*-regulatory elements have been identified. However, *D. melanogaster* males induced to express the sex-determination gene *transformer* (Ferveur et al. 1997) also begin expressing *eloF*, and subsequently produce female-specific dienes with 27 and 29 carbons (Chertemps et al. 2007). *eloF* is thus implicated in *D. melanogaster* diene elongation, and further investigation of genes in the sex-determination pathway, such as DSX, may reveal more about the regulation of *eloF* expression.

desatF and eloF are implicated in CHC production differences between D. simulans and D. sechellia

Quantitative Trait Loci (QTL) analysis implicates *desatF* and *eloF*, among other desaturases and elongases, in the differences in female CHCs between *D. simulans* and *D. sechellia* (Gleason et al. 2005; Gleason et al. 2009). Three major QTL affect the abundance of multiple CHCs differing between *D. simulans* and *D. sechellia* (Figure 1.3). One QTL on the 3rd chromosome affects the abundance of 7,11-HD and is coincident with *desatF*. Another QTL on the 3rd chromosome affects the abundance of 7-T and 7,11-HD and is coincident with *eloF*. This QTL encompasses two clusters of elongases, one of which includes *eloF*. Ten total elongases are found within the two clusters. This gives rise to the possibility that elongases other than *eloF* could be involved in CHC production. Another QTL on the X chromosome is associated with the

quantity of 7-T. This indicates that at least one gene on the X chromosome is affecting CHC production, yet no good candidate genes have been identified for this QTL (Gleason et al. 2009). The 3rd chromosome QTLs that overlap *desatF* and *eloF* act epistatically (Gleason et al. 2009). If *desatF* and *eloF* are involved in sex-specific CHC differences between *D. simulans* and *D. sechellia*, then the epistatic interaction suggests the necessity of both desaturases and elongases in CHC production, if the causative loci are *desatF* and *eloF*.

While expression of both *desatF* and *eloF* may be necessary for 7,11-HD production, expression of those two genes alone may not be sufficient. A study by Hackett (2011) used recombinant inbred lines with small regions of the *D. sechellia* genome in a mostly *D. simulans* genetic background to test the allelic effects of *desatF* and *eloF* on 7,11-HD production. Hybrid individuals heterozygous for the *D. sechellia* and *D. simulans* alleles at either gene produced very little 7,11-HD, although the presence of *D. sechellia* alleles for each gene alters CHCs in the manner predicted by their function (dienes for *desatF* and long carbon chains for *eloF;* Hackett 2011). In the few individuals obtained that were homozygous for both *desatF* and *eloF* alleles from *D. sechellia,* none had a ratio of 7,11-HD to 7-T approaching the ratio seen in *D. sechellia* and were instead closer to the ratio seen in hybrids between *D. simulans* and *D. sechellia*. Thus, *desatF* and *eloF* expression may be necessary for the production of some 7,11-HD, but additional loci may be involved in producing 7,11-HD at the levels produced in *D. sechellia*.

The Hackett (2011) study had limitations. Much of the genome around *desatF* and *eloF* also came from *D. sechellia.* Additionally, because these lines were recombinant inbred lines, the *D. sechellia* genome was not limited to the 3rd chromosome; other chromosomes also included some scattered pieces of the *D. sechellia* genome. Finally, only a small number of females

homozygous for the *D. sechellia* alleles of both *desatF* and *eloF* were measured. However, even with these limitations, these findings support the examination of other elongases and desaturases identified in the QTL study (Gleason et al. 2009).

Research Goals

The primary goal of this research is to test the hypothesis that differential regulation of the genes *desatF* and *eloF* affects differences in CHC production between *D. simulans* and *D. sechellia*. This goal is addressed through the following experiments.

Chapter 2: Gene Expression Patterns Associated with Sex-specific Pheromone Production in D. simulans and D. sechellia.

Expression patterns of candidate elongase and desaturase genes (Gleason et al. 2009) were used to determine if expression was consistent with involvement in female-specific pheromone production. The goal was to determine if any candidate elongases and/or desaturases fit predicted expression patterns, which were 1) Expression of the gene in females of dimorphic species, and no expression in males of dimorphic species or either sex of monomorphic species. 2) No expression of the gene in females of dimorphic species, and expression in males of dimorphic species as well as in both sexes of monomorphic species. Any gene that matched one of these patterns was considered for further investigation of potential involvement in CHC production, and any gene whose expression did not match one of the patterns was ruled out as a candidate gene. Expression of each candidate gene was measured in abdominal cuticle tissue from two monomorphic species (*D. simulans* and *D. mauritiana*), and three dimorphic species (*D. sechellia, D. melanogaster*, and *D. erecta*). Only *desatF* and *eloF* were found to fit the

pattern expected from the hypothesis. *desatF* and *eloF* were then further analyzed in Chapter 3 of this study for their effects on CHC production and courtship in *D. simulans* and *D. sechellia*.

The second goal of this research was to test the hypothesis that *cis*-regulatory differences are causing differential expression of *desatF* and *eloF* between *D. simulans* and *D. sechellia*. This goal was addressed by measuring the allele-specific mRNA expression of *desatF* and *eloF* in hybrids of *D. simulans* and *D. sechellia.* I predicted that expression of both genes would come only from the *D. sechellia* allele, due to a *cis*-acting regulatory difference between the two species. The results supported the hypothesis that *cis*-regulation is a causative factor in expression differences between *D. simulans* and *D. sechellia* females.

Chapter 3: Introgression of desatF and eloF Alleles Affects Gene Expression and Courtship between D. simulans and D. sechellia.

The goal of this research was to test the hypothesis that *D. simulans* alleles of *desatF* and *eloF* in a *D. sechellia* genetic background will alter *desatF* and *eloF* gene expression and courtship behavior. *D. simulans* alleles of *desatF* and *eloF* were introgressed into a *D. sechellia* genetic background through the production of near isogenic lines. Females of lines homozygous for the *D. simulans* allele of *desatF* were predicted to lack expression of *desatF*. Females of lines containing the *D. simulans* allele of *eloF* were predicted to lack expression of *eloF*. Females of a line containing the *D. simulans* allele of both *desatF* and *eloF* were predicted to lack expression of both *desatF* and *eloF*. Females of introgression lines were also predicted to be courted by *D. simulans* males, despite the usual lack of courtship from *D. simulans* males toward *D. sechellia* females.

Conclusion

The roles of desaturases and elongases in *D. melanogaster* pheromone biosynthesis make these enzymes attractive candidates for study in other species within the *D. melanogaster* group. In particular, the sex-specific expression of *desatF* and *eloF*, and their connections with sex-specific CHC production, suggests these genes are influencing reproductive isolation in species that use CHCs in sexual signaling. The study of the genetic basis of CHC differences between *D. simulans* and *D. sechellia* further characterizes the roles of *desatF* and *eloF* in maintaining a reproductive barrier between two closely related species. This dissertation research seeks to investigate a genetic cause behind the reproductive isolation between *D. simulans* and *D. sechellia* by 1) Testing the expression of candidate elongases and desaturases to determine if any are expressed in a pattern suggesting they are involved CHC differences between females of *D. simulans* and *D. sechellia*, 2) Testing whether expression differences of *desatF* and *eloF* are consistent with *cis*-regulatory effects, and 3) Testing whether the introgression of alleles of *desatF* and/or *eloF* from one species into another alters expression of the genes and affects courtship behavior. Future research will be needed to determine the specific functional role of *desatF* and *eloF* in CHC biosynthesis in *D. simulans* and *D. sechellia*, as well as further characterization of the full CHC biosynthesis pathway, which may include genes not described in this dissertation.

Literature Cited

Attrill, H., K. Falls, J. L. Goodman, G. H. Millburn, G. Antonazzo, A. J. Rey, and S. J. Marygold. 2016. FlyBase: establishing a gene group resource for *Drosophila melanogaster*. Nucleic Acids Research 44:D786-792.

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

Borneman, A. R., Z. D. Zhang, J. Rozowsky, M. R. Seringhaus, M. Gerstein, and M. Snyder. 2007. Transcription factor binding site identification in yeast: a comparison of high-density oligonucleotide and PCR-based microarray platforms. Functional Integrative Genomics 7:335-345.

Bousquet, F., T. Nojima, B. Houot, I. Chauvel, S. Chaudy, S. Dupas, D. Yamamoto, and J. F. Ferveur. 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 of the United States of America 109:249-254.

Bradley, R. K., X. Y. Li, C. Trapnell, S. Davidson, L. Pachter, H. C. Chu, L. A. Tonkin, M. D. Biggin, and M. B. Eisen. 2010. Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related *Drosophila* species. PLoS Biology 8:e1000343.

Bustamante, C. D., A. Fledel-Alon, S. Williamson, R. Nielsen, M. T. Hubisz, S. Glanowski, D. M. Tanenbaum, T. J. White, J. J. Sninsky, R. D. Hernandez, D. Civello, M. D. Adams, M. Cargill, and A. G. Clark. 2005. Natural selection on protein-coding genes in the human genome. Nature 437:1153-1157.

Carroll, S. B. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell 134:25-36.

Cheatle Jarvela, A. M. and V. F. Hinman. 2015. Evolution of transcription factor function as a mechanism for changing metazoan developmental gene regulatory networks. Evodevo 6:3.

Chertemps, T., L. Duportets, C. Labeur, R. Ueda, K. Takahashi, K. Saigo, and C. Wicker-Thomas. 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.

Chertemps, T., L. Duportets, C. Labeur, M. Ueyama, and C. Wicker-Thomas. 2006. A female-specific desaturase gene responsible for diene hydrocarbon biosynthesis and courtship behaviour in *Drosophila melanogaster*. Insect Molecular Biology 15:465-473.

Chertemps, T., L. Duportets, C. Labeur, and C. Wicker-Thomas. 2005. A new elongase selectively expressed in *Drosophila* male reproductive system. Biochemical and Biophysical Research Communications 333:1066-1072.

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

Coolon, J. D., C. J. McManus, K. R. Stevenson, B. R. Graveley, and P. J. Wittkopp. 2014. Tempo and mode of regulatory evolution in *Drosophila*. Genome Researchearch 24:797-808.

Coolon, J. D., W. Webb, and P. J. Wittkopp. 2013. Sex-specific effects of cis-regulatory variants in *Drosophila melanogaster*. Genetics 195:1419-1422.

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

Dallerac, R., C. Labeur, J. M. Jallon, D. C. Knippie, W. L. Roelofs, and C. Wicker-Thomas. 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 of the United States of America 97:9449-9454.

Dobzhansky, T. 1935. A critique of the species concept in biology. Philosophy of Science 2:344-345.

Emerson, J. J., L. C. Hsieh, H. M. Sung, T. Y. Wang, C. J. Huang, H. H. Lu, M. Y. Lu, S. H. Wu, and W. H. Li. 2010. Natural selection on cis and trans regulation in yeasts. Genome Research 20:826-836.

Etges, W. J. 2014. No boundaries: genomes, organisms, and ecological interactions responsible for divergence and reproductive isolation. Journal of Heredity 105 Suppl 1:756-770.

Fang, S., C. T. Ting, C. R. Lee, K. H. Chu, C. C. Wang, and S. C. Tsaur. 2009. Molecular evolution and functional diversification of fatty acid desaturases after recurrent gene duplication in *Drosophila*. Molecular Biology and Evolution 26:1447-1456.

Fang, S., R. Yukilevich, Y. Chen, D. A. Turissini, K. Zeng, I. A. Boussy, and C. I. Wu. 2012. Incompatibility and competitive exclusion of genomic segments between sibling *Drosophila* species. PLoS Genetics 8.

Ferveur, J. F. 1997. The pheromonal role of cuticular hydrocarbons in *Drosophila melanogaster*. Bioessays 19:353-358.

Ferveur, J. F. 2005. Cuticular hydrocarbons: their evolution and roles in *Drosophila* pheromonal communciation. Behavioral Genetics 35:279-295.

Ferveur, J. F., F. Savarit, C. J. O'Kane, G. Sureau, R. J. Greenspan, and J. M. Jallon. 1997. Genetic feminization of pheromones and its behavioral consequences in *Drosophila* males. Science 276:1555-1558.

Frise, E., A. S. Hammonds, and S. E. Celniker. 2010. Systematic image-driven analysis of the spatial *Drosophila* embryonic expression landscape. Molecular Systems Biology 6:345.

Garfield, D. A. and G. A. Wray. 2010. The Evolution of Gene Regulatory Interactions. Bioscience 60:15-23.

Garrigan, D., S. B. Kingan, A. J. Geneva, P. Andolfatto, A. G. Clark, K. R. Thornton, and D. C. Presgraves. 2012. Genome sequencing reveals complex speciation in the *Drosophila simulans* clade. Genome Research 22:1499-1511.

Gleason, J. M., J. M. Jallon, J. D. Rouault, and M. G. Ritchie. 2005. Quantitative trait loci for cuticular hydrocarbons associated with sexual isolation between *Drosophila simulans* and *D. sechellia*. Genetics 171:1789-1798.

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

Gordon, K. L. and I. Ruvinsky. 2012. Tempo and mode in evolution of transcriptional regulation. PLoS Genetics 8:e1002432.

Gruber, J. D., K. Vogel, G. Kalay, and P. J. Wittkopp. 2012. Contrasting properties of gene-specific regulatory, coding, and copy number mutations in *Saccharomyces cerevisiae*: frequency, effects, and dominance. PLoS Genetics 8:e1002497.

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

Jallon, J. M. and J. R. David. 1987. Variations in cuticular hydrocarbons among the eight species of the *Drosophila Melanogaster* subgroup. Evolution 41:294-302.

Keays, M. C., D. Barker, C. Wicker-Thomas, and M. G. Ritchie. 2011. Signatures of selection and sex-specific expression variation of a novel duplicate during the evolution of the *Drosophila* desaturase gene family. Molecular Ecology 20:3617-3630.

Kohler, K., E. Brunner, X. L. Guan, K. Boucke, U. F. Greber, S. Mohanty, J. M. Barth, M. R. Wenk, and E. Hafen. 2009. A combined proteomic and genetic analysis identifies a role for the lipid desaturase Desat1 in starvation-induced autophagy in *Drosophila*. Autophagy 5:980-990.

Kyriacou, C. P. and J. C. Hall. 1982. The function of courtship song rhythms in *Drosophila*. Animal Behavior 30:794-801.

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

Landry, C. R., P. J. Wittkopp, C. H. Taubes, J. M. Ranz, A. G. Clark, and D. L. Hartl. 2005. Compensatory cis-trans evolution and the dysregulation of gene expression in interspecific hybrids of *Drosophila*. Genetics 171:1813-1822.

Legendre, A., X. X. Miao, J. L. Da Lage, and C. Wicker-Thomas. 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.

Lynch, V. J., G. May, and G. P. Wagner. 2011. Regulatory evolution through divergence of a phosphoswitch in the transcription factor CEBPB. Nature 480:383-386.

Mackay, T. F. 2004. The genetic architecture of quantitative traits: lessons from *Drosophila*. Current Opinion in Genetics & Development 14:253-257.

McManus, C. J., J. D. Coolon, M. O. Duff, J. Eipper-Mains, B. R. Graveley, and P. J. Wittkopp. 2010. Regulatory divergence in *Drosophila* revealed by mRNA-seq. Genome Research 20:816-825.

Meiklejohn, C. D., J. D. Coolon, D. L. Hartl, and P. J. Wittkopp. 2014. The roles of cis- and trans-regulation in the evolution of regulatory incompatibilities and sexually dimorphic gene expression. Genome Research 24:84-95.

Molodtsova, D., B. A. Harpur, C. F. Kent, K. Seevananthan, and A. Zayed. 2014. Pleiotropy constrains the evolution of protein but not regulatory sequences in a transcription regulatory network influencing complex social behaviors. Frontiers in Genetics 5:431.

Montell, C. 2009. A taste of the *Drosophila* gustatory receptors. Current Opinion in Neurobiology 19:345-353.

Musselman, L. P., J. L. Fink, P. V. Ramachandran, B. W. Patterson, A. L. Okunade, E. Maier, M. R. Brent, J. Turk, and T. J. Baranski. 2013. Role of fat body lipogenesis in protection against the effects of caloric overload in *Drosophila*. Journal of Biological Chemistry 288:8028-8042.

Ng, W. C., J. S. Chin, K. J. Tan, and J. Y. Yew. 2015. The fatty acid elongase *Bond* is essential for *Drosophila* sex pheromone synthesis and male fertility. Nature Communications 6:8263.

Orr, H. A., J. P. Masly, and D. C. Presgraves. 2004. Speciation genes. Current Opinion in Genetics & Development 14:675-679.

Parisi, F., S. Riccardo, S. Zola, C. Lora, D. Grifoni, L. M. Brown, and P. Bellosta. 2013. dMyc expression in the fat body affects DILP2 release and increases the expression of the fat desaturase Desat1 resulting in organismal growth. Developmental Biology 379:64-75.

Ritchie, M. G., R. M. Townhill, and A. Hoikkala. 1998. Female preference for fly song: playback experiments confirm the targets of sexual selection. Animal Behavior 56:713-717.

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

Sawamura, K., J. Roote, C. I. Wu, and M. T. Yamamoto. 2004. Genetic complexity underlying hybrid male sterility in *Drosophila*. Genetics 166:789-796.

Shirangi, T. R., H. D. Dufour, T. M. Williams, and S. B. Carroll. 2009. Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*. PLoS Biology 7 (8): e1000168.

Smadja, C. and R. K. Butlin. 2009. On the scent of speciation: the chemosensory system and its role in premating isolation. Heredity (Edinb) 102:77-97.

Spieth, H. T. 1974. Courtship behavior in *Drosophila*. Annual Review of Entomology 19:385-405.

Stern, D. L. 2000. Evolutionary developmental biology and the problem of variation. Evolution 54:1079-1091.

Stocker, R. F. and N. Gendre. 1989. Courtship behavior of *Drosophila* genetically or surgically deprived of basiconic sensilla. Behavioral Genetics 19:371-385.

Tauber, E. and D. F. Eberl. 2002. The effect of male competition on the courtship song of *Drosophila melanogaster*. Journal of Insect Behavior 15:109-120.

Ting, C. T., S. C. Tsaur, M. L. Wu, and C. I. Wu. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. Science 282:1501-1504.

Tirosh, I., S. Reikhav, A. A. Levy, and N. Barkai. 2009. A yeast hybrid provides insight into the evolution of gene expression regulation. Science 324:659-662.

Tomancak, P., A. Beaton, R. Weiszmann, E. Kwan, S. Shu, S. E. Lewis, S. Richards, M. Ashburner, V. Hartenstein, S. E. Celniker, and G. M. Rubin. 2002. Systematic determination of patterns of gene expression during *Drosophila* embryogenesis. Genome Biology 3:RESEARCH0088.

Tomaru, M., H. Matsubayashi, and Y. Oguma. 1998. Effects of courtship song in interspecific crosses among the species of the *Drosophila auraria* complex (*Diptera: Drosophilidae*). Journal of Insect Behavior 11:383-398.

Wagner, G. P. and V. J. Lynch. 2008. The gene regulatory logic of transcription factor evolution. Trends in Ecology and Evolution 23:377-385.

Weirauch, M. T., A. Yang, M. Albu, A. G. Cote, A. Montenegro-Montero, P. Drewe, H. S. Najafabadi, S. A. Lambert, I. Mann, K. Cook, H. Zheng, A. Goity, H. van Bakel, J. C. Lozano, M. Galli, M. G. Lewsey, E. Huang, T. Mukherjee, X. Chen, J. S. Reece-Hoyes, S. Govindarajan, G. Shaulsky, A. J. Walhout, F. Y. Bouget, G. Ratsch, L. F. Larrondo, J. R. Ecker, and T. R. Hughes. 2014. Determination and inference of eukaryotic transcription factor sequence specificity. Cell 158:1431-1443.

Weiszmann, R., A. S. Hammonds, and S. E. Celniker. 2009. Determination of gene expression patterns using high-throughput RNA in situ hybridization to whole-mount *Drosophila* embryos. Nature Protocols 4:605-618.

Wittkopp, P. J., B. K. Haerum, and A. G. Clark. 2004. Evolutionary changes in cis and trans gene regulation. Nature 430:85-88.

Wittkopp, P. J., B. K. Haerum, and A. G. Clark. 2008a. Independent effects of cis- and trans-regulatory variation on gene expression in *Drosophila melanogaster*. Genetics 178:1831-1835.

Wittkopp, P. J., B. K. Haerum, and A. G. Clark. 2008b. Regulatory changes underlying expression differences within and between *Drosophila* species. Nature Genetics 40:346-350.

Wittkopp, P. J. and G. Kalay. 2012. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nature Reviews Genetics 13:59-69.

Wray, G. A. 2007. The evolutionary significance of cis-regulatory mutations. Nature Reviews Genetics 8:206-216.

Yvert, G., R. B. Brem, J. Whittle, J. M. Akey, E. Foss, E. N. Smith, R. Mackelprang, and L. Kruglyak. 2003. Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. Nature Genetics 35:57-64.

Zhang, X., A. J. Cal, and J. O. Borevitz. 2011. Genetic architecture of regulatory variation in *Arabidopsis thaliana*. Genome Research 21:725-733.

Figure 1.1

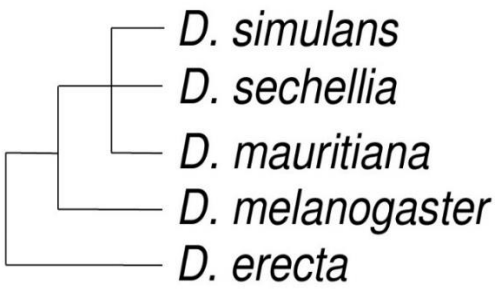
D. simulans
D. sechellia
D. mauritiana
D. melanogaster
D. erecta

Figure 1.1. Phylogeny of the focal species in the *Drosophila melanogaster* subgroup (Adapted from Garrigan et al. 2012).

Figure 1.2

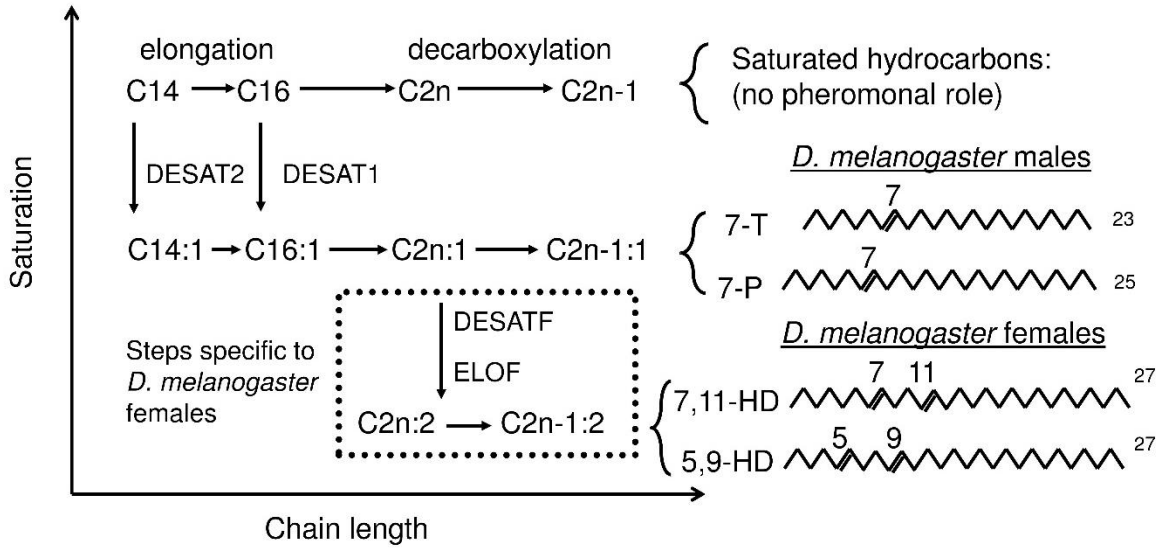
elongation
decarboxylation
C14
C16
C2n
C2n-1
Saturated hydrocarbons:
(no pheromonal role)
DESAT2
DESAT1
D. melanogaster males
7
7-T
23
7
7-P
25
C14:1
C16:1
C2n:1
C2n-1:1
DESATF
ELOF
Steps specific to
D. melanogaster
females
C2n:2
C2n-1:2
D. melanogaster females
7
11
7,11-HD
27
5
9
5,9-HD
27
Saturation
Chain length

Figure 1.2. **Biosynthesis of primary cuticular hydrocarbons in *D. melanogaster*.** DESAT1 and DESAT2 act in both males and females to form a single double bond at either the 5^{th} or 7^{th} carbon, depending on the strain. In *D. melanogaster* females, the carbon chain is further desaturated and elongated by a female-specific desaturase (DESATF) and elongase (ELOF), respectively. Adapted from Legendre et al. (2008).

Figure 1.3

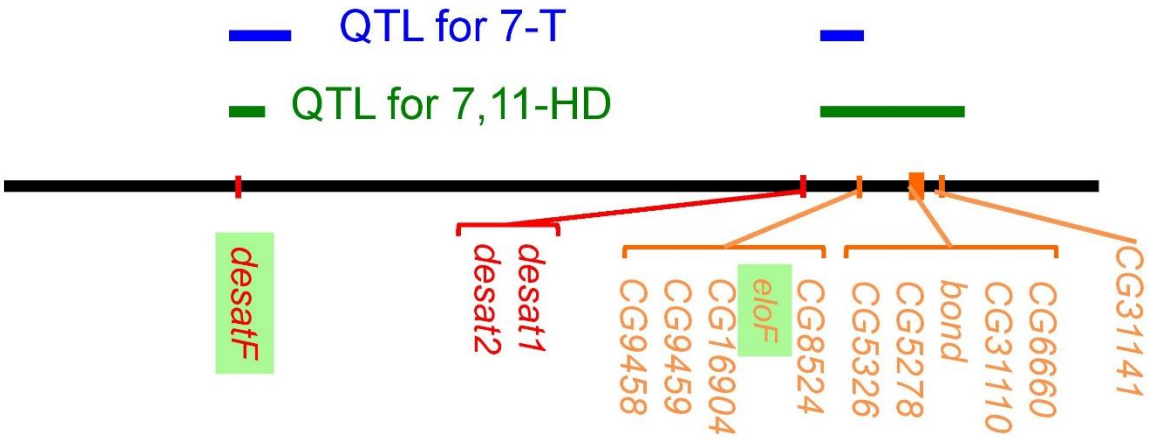
QTL for 7-T
QTL for 7,11-HD
desatF
desat1
desat2
CG9458
CG9459
CG16904
eloF
CG8524
CG5326
CG5278
bond
CG31110
CG6660
CG31141

Figure 1.3. Quantitative trait loci (QTL) mapping of 7-T and 7,11-HD differing between *D. simulans* and *D. sechellia* females. Two QTL regions are shown on the 3rd chromosome. QTL for 7-T are in blue, and QTL for 7,11-HD are in green. A QTL on the X chromosome is not shown. Candidate desaturases are in red, and candidate elongases are in orange. Adapted from Gleason et al. (2009).

Chapter 2

Gene Expression Patterns Associated with Sex-Specific Pheromone Production in *Drosophila simulans and D. sechellia*

Introduction

The study of the expression of genes that affect mating traits is important in understanding how closely related species can be kept reproductively isolated. Alterations to gene pathways can induce chemical and behavioral changes that may influence population divergence, particularly when gene pathway modifications change chemical cues involved in courtship (Smadja and Butlin 2009). Species- and sex-specific expression differences of genes involved in courtship trait biosynthesis pathways can affect barriers to reproduction (Chertemps et al. 2007; Legendre et al. 2008). In *Drosophila*, chemosensory signaling is a highly variable courtship trait among species (Cobb and Jallon 1990), and changes in the regulation of gene pathways affecting chemosensory signaling can rapidly lead to reproductive isolation (Shirangi et al. 2009). Such changes can cause reproductive isolation by leading to species-specific phenotypes (Mackay 2004; Wittkopp and Kalay 2012). Therefore, the study of how changes in gene expression alter biosynthesis pathways of traits affecting chemosensory signaling and mating behavior traits can be informative in understanding reproductive isolation and, ultimately, speciation (Smadja and Butlin 2009).

Biosynthesis pathways consist of a succession of different protein actions that work together in a particular order, like an assembly line, to build a biological product. The presence of the proteins involved in a given pathway is dependent on the proper time- and tissue-specific expression of the genes encoding those proteins. The lack of a given protein needed at a particular point in a pathway can alter the pathway's product, as can the addition of new proteins playing new roles in the production line. Changes in gene expression affecting proteins in

biosynthesis pathways can thus lead to novel biological products (Selifonova et al. 2001; Wagner and Lynch 2010).

In *Drosophila melanogaster*, the biosynthesis pathway leading to the production cuticular hydrocarbons (CHCs), which act as pheromones, is known to have undergone sex- and species-specific changes through alterations in gene expression (Chertemps et al. 2007; Legendre et al. 2008). Biosynthesis pathways that produce CHCs differing between females of different species can influence courtship barriers, due to male preference for particular CHC types. This study examines the expression of genes that are potentially involved in differential pheromone biosynthesis pathways between species of the *D. melanogaster* subgroup, in order to identify specific genes involved in the maintenance of reproductive isolation.

Cuticular hydrocarbons and Drosophila courtship

As discussed in Chapter 1, *Drosophila* courtship is influenced by the presence of cuticular hydrocarbons (CHCs) secreted onto the cuticle of the fly (Ferveur 1997). A *Drosophila* male detects CHCs by tapping the abdomen of a potential mate with his foreleg, which contains chemosensory receptors (Jallon 1984). This behavior enables both sex- and species-recognition based on the CHC profile of the potential mate (Billeter et al. 2009), and further male courtship depends on the male's attraction or aversion to the female's CHCs.

CHC profiles species within the *D. melanogaster* subgroup can be monomorphic or dimorphic with respect to males and females. *D. simulans* females produce predominantly 7-T while *D. sechellia* females produce predominantly 7,11-HD (Cobb and Jallon 1990). *D. melanogaster* CHCs are the same as *D. sechellia* (7-T in males, 7,11-HD in females), and *D. erecta* has the same general trend of longer, less saturated CHCs in females than in males. The predominance of 7-T in both sexes occurs in *D. mauritiana*. The relationship (Figure 1.1) among

the three sibling species, *D. simulans*, *D. sechellia* and *D. mauritiana*, is not well established (Kastanis et al. 2003; Garrigan et al. 2012a). It is not clear if *D. mauritiana* and *D. simulans* both lost sexual dimorphism or if monomorphism was the ancestral state and dimorphism was gained in *D. sechellia.*

Candidate genes for CHC production differences between D. simulans and D. sechellia

As discussed in chapter 1, Quantitative Trait Loci (QTL) analysis implicated three major QTL in the differences in female CHCs between *D. simulans* and *D. sechellia* (Gleason et al. 2005; Gleason et al. 2009). In association with two of the QTL regions, I identified *desatF* and two other desaturases, as well as *eloF* and ten other elongases. I hypothesized that the differential expression of *desatF* and *eloF* affects CHC production differences between *D. simulans* and *D. sechellia*. As described in chapter 1, expression of both *desatF* and *eloF* may be necessary for 7,11-HD production, but expression of those two genes alone may not be sufficient (Hackett 2011). While the QTL analysis implicated regions containing *desatF* and *eloF*, there were several other desaturases and elongases within or near the QTL regions. Therefore, I sought to test if any of the candidate desaturases and/or elongases had a pattern of expression consistent with sex-specific CHC production. I tested the expression of the candidate genes in two monomorphic species (*D. mauritiana* and *D. simulans*), and three dimorphic species (*D. sechellia*, *D. melanogaster*, and *D. erecta*). Because 7,11-HD is female-specific, my prediction was that genes showing female-specific gene expression in dimorphic species, and no gene expression in monomorphic species, are likely to be involved in CHC production differences between *D. simulans* and *D. sechellia*. Alternatively, gene expression could be male-specific in dimorphic species, and be present in both sexes of monomorphic species.

Determination of cis- versus trans- regulation in expression of desatF and eloF

As described in chapter 1, *cis*- regulation of gene expression involves regulatory elements affecting the chromosome on which those elements are located, such as promoters and enhancers. *Trans*-regulation involves regulatory proteins such as transcription factors that can move from one chromosome to another to activate or repress gene expression. Because *cis*-regulation is thought to have a large influence on divergent gene expression (Wittkopp et al. 2004), I determined if differences in *desatF* and *eloF* expression between *D. simulans* and *D. sechellia* females was due to differences in *cis*- elements or *trans*- elements. I hypothesized that *cis*- regulatory differences are responsible for differential expression of *desatF* and *eloF* between *D. simulans* and *D. sechellia*. To test this hypothesis, I examined allele-specific expression of both genes in hybrid females. I predicted that expression of each gene would come only from the *D. sechellia* allele, due to a *cis*-regulatory difference in the *D. simulan*s allele that prevents expression, regardless of any *trans*-acting protein from the *D. sechellia* allele attempting to activate expression of the *D. simulans* allele.

Materials and Methods

Fly stocks and culturing

Cultures of each species, *D. sechellia* strain David 4A (described in Gleason and Ritchie 2004), *D. simulans* strain f^2*;nt,pm,e (*described in Gleason and Ritchie 2004*), D. melanogaster* strain 14021-0231.36, *D. mauritiana* strain 14021-0241.01, and *D. erecta* strain 14021-0224.01 (all three from the Drosophila Species Stock Center, UC San Diego) were maintained on standard cornmeal-molasses food in 25 x 95 mm polystyrene vials in populations of roughly twenty flies.

The flies were kept at 25°C on a 12 hour light/12 hour dark cycle. Hybrid flies were made by crossing ten *D. simulans* virgin females with ten *D. sechellia* males.

Tissue dissection

For each species, as well as hybrids, virgin males and virgin females were collected separately within four hours of eclosion, separated by sex, and kept in groups of 10 in vials. At four days post-eclosion, when flies are producing adult cuticular hydrocarbons, the flies were briefly anesthetized under CO_2 and dissected in RNAlater (Qiagen) on a glass microscope slide. For each preparation, fifteen abdomens were removed with forceps and the internal tissues, including testes, accessory glands or ovaries, were discarded. The abdominal cuticle was submerged in RNAlater in 1.5 ml Eppendorf tubes, and stored at -20°C.

RNA extraction and cDNA synthesis

For RNA extraction, the tissue preparations were removed from the RNAlater and placed in lysis buffer (Purelink® RNA Minikit, Life Technologies). The tissue was ground with a plastic tissue grinder in a 1.5 ml microcentrifuge tube until thoroughly disrupted. The RNA was extracted following the protocol of the Purelink® RNA Minikit (Life Technologies). Genomic DNA was removed from each preparation using Turbo DNA-free™ kit (Life Technologies). DNase was removed from each preparation with the DNase Inactiviation Reagent (Life Technologies). cDNA was synthesized from each RNA preparation using the iScript™ cDNA synthesis kit (Bio-Rad). Lack of (Table 1) amplification of tissue-specific ovary (*Femcoat*) and testes (*Acp26*) transcripts was used to confirm the removal of gonadal tissues before progressing with further analysis.

Analysis of gene expression by qPCR

qPCR was performed using the SYBR® Select Master Mix (Life Technologies). Primers (Table 2.1) were designed for each gene to amplify a segment of DNA less than 200 bp in length. All primers were designed to flank each side of an intron, with the exception of those for *desatF*, which lacks introns. *rp49* was used as a positive control and as a calibrator for the relative quantification method. An NTC (no template control) was used as a negative control for each sample to check for DNA contamination in the PCR mix. Amplification was performed in a StepOnePlus® Real Time PCR System (Life Technologies) on two biological replicates for each species and sex with two technical replicates per biological replicate.

Analysis was performed using StepOnePlus® Real Time PCR System software (Life Technologies). Cycle threshold (CT) values for each gene were determined for each sample. Primer efficiency (E) for each locus was determined using DART-PCR version 1.0 (Peirson et al. 2003). Normalization of expression amounts to the ribosomal protein gene *rp49* allowed for the comparison of gene expression across different samples. Relative expression values were determined using the Delta CT method (Hellemans et al. 2007) and described as follows: The R_0 value (the amount of starting DNA material) was determined for each sample using the formula $R_0 = (E+1)^{-CT}$. Two technical replicates per biological replicate were averaged to get mean biological R_0 values. Mean R_0 values for biological replicates were divided by mean R_0 values for the corresponding *rp49* samples. The mean of normalized R_0 values for biological replicates within each sex and species was used as the value to represent expression amount. I used Microsoft Excel software to perform Student t-tests to determine differences in expression between males and females for each species.

qPCR of D. simulans/D. sechellia hybrids

To determine which allele was expressed in hybrids of *D. simulans* and *D. sechellia*, I performed allele specific quantitative PCR on hybrids and compared them to *D. simulans* and *D. sechellia* females. Tissue dissections, RNA extractions, and cDNA synthesis were performed as described above. I designed Taqman probes (Life Technologies) with at least one base pair difference specific for either the *D. simulans* or the *D. sechellia* allele of *desatF* and *eloF* (Table 2.2). The specificity of the probes for each allele was confirmed with genomic DNA. I tested a range of annealing and extension temperatures to optimize probe specificity for each allele. I found the optimal setting for distinguishing between *D. simulans* and *D. sechellia eloF* alleles was 40 cycles of 15 seconds annealing at 58 ˚C, and 30 seconds extension at 74˚C. For *desatF*, the optimal setting was 40 cycles of 15 seconds annealing at 60˚C, and 30 seconds extension at 74 ˚C. I used the same amplification primers for *desatF* and *eloF* as described above for the qPCR assay. A no-template control was used for each sample to check for DNA contamination in the qPCR mix.

Amplification was performed on two biological replicates for species and sex with three technical replicates per biological replicate. Analysis was performed using StepOnePlus® Real Time PCR System software (Life Technologies). CT values for each species/allele of *desatF* and *eloF* were normalized to the *D. sechellia* CT value for the respective gene by subtracting the species/allele value from the *D. sechellia* value. The normalized values were then changed to 1/(normalized value), giving the *D. sechellia* expression a value of 1.0. Relative fold difference values were then determined by dividing the *D. sechellia* expression value by the species/allele value. Mean fold expression differences were then reported relative to *D. sechellia* expression.

Results

qPCR results

The expression of each candidate gene was measured in both sexes of all species using quantitative RT-PCR (qPCR; Figure 2.1). Of the desaturase genes I tested, only *desatF* had a trend toward female-specific expression in dimorphic species. Likewise, of the elongase genes I tested, only *eloF* had a trend of female-specific expression in dimorphic species. While only *D. melanogaster* had statistically significant differences by sex, *D. sechellia* had a trend toward female expression, while the monomorphic species had only minimal expression, if any at all. These results suggest *desatF* and *eloF* are the best candidates for influencing CHC production differences between *D. simulans* and *D. sechellia*.

Desaturases

I hypothesized that good candidate genes would have a predictable pattern of expression, which was that dimorphic species would have female sex-biased expression, and monomorphic species would either lack expression in both sexes or have a similar amount of expression in both species. In the dimorphic species *D. sechellia, D. melanogaster,* and *D. erecta,* females had more *desatF* expression than males, however only in *D. melanogaster* was the difference between the sexes statistically significant (t-test, P=0.007). Low expression of *desatF* was detected in males of *D. sechellia*. Although this expression was minimal, the result of the t-test gave a P value of 0.34, which is not a significant difference. The P value for male/female differences in *D. erecta* was also not significant (t-test, P=0.38). The data still have a trend of female expression in dimorphic species and no expression in monomorphic species.

Desat1 had higher expression in males of dimorphic species. However, only the sex difference within *D. sechellia* (t-test, *P*=0.04) was statistically significant. While *desat1* may be playing a role in the abdominal tissue in each species, it appears to be active in both sexes.

Desat2 was expressed in *D. melanogaster* females but not males (t-test, *P*=0.14). Minimal to no expression was found in the both the females and males of *D. simulans*, *D. mauritiana*, and *D. erecta*. *D. sechellia* expressed *desat2* equally among the sexes (t-test, *P*=0.13). *desat2* is possibly playing some sex-specific role in *D. melanogaster* females and *D. sechellia* males. However, the opposite sex specificity between *D. melanogaster* and *D. sechellia*, and the absence of *desat2* in *D. erecta* make *desat2* unlikely to be involved in sex-specific production of 7-T and 7,11-HD.

EloF gene cluster on 3rd chromosome – CG9458, CG9459, eloF, CG16904, CG8534

The *eloF* cluster of elongases centers on *eloF*. All five genes exist within a roughly 8 Kb region on the 3rd chromosome. The gene sequences are all highly similar, and likely represent a region of gene duplications (Howard and Blomquist 2005; Fang et al. 2009). *CG9458* was expressed in both sexes of each species, except *D. mauritiana* males, with no obvious sex-specific pattern. *CG9459* had minimal to no expression in any of the species (data not shown). *CG16904* had male biased expression in *D. sechellia*, *D. melanogaster*, and *D. erecta*, although the sex difference was only statistically significant in *D. melanogaster* (t-test, *P*=0.007). This makes *CG16904* a potentially interesting gene to investigate.

CG8534 had an interesting pattern of expression that is also potentially worth further investigation. Females of *D. sechellia*, *D. melanogaster*, and *D. erecta* had more expression than males, males of *D. simulans* had more expression than females, and neither sex had expression in

D. mauritiana. Despite appearing to match my hypothesis, statistical analysis showed differences to be non-significant.

eloF was the only elongase I tested to be highly expressed in females of *D. sechellia* (t-test, *P*=0.1) and *D. melanogaster* (t-test, *P*=0.008) compared to expression in males of those species. While *D. erecta* had minimal female expression; the full *eloF* gene is not present in the *D. erecta* genome. The qPCR primers bind to the fragment that is still present, so the observed expression could have been due to the presence of RNA containing this fragment. *D. mauritiana* had minimal to no expression in either sex. In the monomorphic species *D. simulans*, an extremely low level of expression of *eloF* was detected in males, but the sex-difference was not significant. Thus, among the *eloF* cluster, only the expression pattern of *eloF* suggests involvement in CHC production differences between monomorphic and dimorphic species.

bond gene cluster on 3rd chromosome – CG5326, CG5278, bond, CG33110, CG6660

The *bond* cluster lies within a 391Kb stretch on the 3rd chromosome. *CG5326* was expressed with a female-biased pattern in *D. mauritiana* (t-test, *P*=0.005), but had no sex difference in *D. simulans*. Dimorphic species had more expression in males, with the differences being significant in *D. sechellia* (t-test, *P*=0.048) and *D. melanogaster* (t-test, *P*=0.011). *CG5278* was highly expressed in *D. erecta* males, but was only minimally expressed, if at all, in the other species. *Bond* had a sex-specific expression difference only in *D. melanogaster*, with males having higher expression than females, although the difference was not statistically significant (t-test, *P*=0.11). *D. sechellia* had slightly higher *bond* expression in males, while *D. erecta* had slightly higher expression of *bond* in females, but the difference was not significant. *D. simulans* and *D. mauritiana* had no difference in *bond* expression between the sexes.

CG33110 was most highly expressed in *D. melanogaster* and *D. erecta*, but sex differences were not significant. Minimal expression of *CG33110* was found in either sex of *D. simulans*, *D. mauritiana*, and *D. sechellia*. Minimal expression of *CG6660* was detected in both sexes of all species, with the exception of *D. erecta*, wherein higher female expression was observed. However, no significant sex differences were found. Again, these patterns for each of the genes in the *bond* cluster deviate from the expectation of female specificity in only dimorphic species. Thus, these genes are not likely involved in the difference in 7-T and 7,11-HD production in the species tested.

CG31141

CG31141 is located beyond the *bond* cluster farther out on the chromosome. Expression of C*G31141* was mostly limited to *D. erecta* males, with minimal to no expression in each of the other species. This *D. erecta*-specific expression pattern makes this gene also unlikely to be involved in in the difference in 7-T and 7,11-HD production in the species tested.

Expression of desatF and eloF in hybrids

I tested allele-specific expression of *desatF* and *eloF* in hybrid females made with *D. simulans* females and *D. sechellia* males (Figure 2.2). All primers were confirmed to work on genomic DNA prior to use in qPCR. *desatF* expression was detected from the *D. sechellia* allele in hybrids, while only a very minimal amount was detected from the *D. simulans* allele. Similarly, *eloF* expression was detected from the *D. sechellia* allele in hybrids, while no expression was detected from the *D. simulans* allele. Hybrid expression of the *D. sechellia eloF* allele was lower than the expression of *eloF* in the *D. sechellia* strain females. These results

suggest a *cis*-regulatory difference in expression of *desatF* and *eloF* between *D. simulans* and *D. sechellia* females.

Discussion

To understand how novel traits arise through evolution, it is useful to identify trait-specific gene expression (Williams et al. 2008). Differences in the regulation of gene expression between species are known to be a cause of novel morphological traits (Wagner and Lynch 2010). Novel, sexually dimorphic traits, such as the male-specific abdominal pigmentation of *D. melanogaster,* have evolved through alterations to genetic regulatory pathways (Williams et al. 2008). Other examples of sexually dimorphic traits controlled by sex-specific gene expression are the male "sword" of the swordtail fish (Zauner et al. 2003) and *Scr* in *Drosophila*, which has a male-specific expression pattern involved with sex comb morphology (Graze et al. 2007). Thus, sex-specific gene regulation is important in the expression of dimorphic traits.

By examining expression patterns of desaturases and elongases identified in a QTL study (Gleason et al. 2009), I have begun to determine what genes may be causing differential CHC production between females of *D. simulans* and *D. sechellia*. Working with my hypothesis that genes involved in sex-specific CHC production will be expressed only in females of dimorphic species, I have found that the candidate genes *desatF* and *eloF* fit the expected pattern of expression. Of all elongases tested in this study, *eloF* is the only gene to have female-specific expression in dimorphic species while lacking expression in both sexes of monomorphic species.

Although *D. erecta* is a dimorphic species, analysis of the *D. erecta* genome has revealed that the full *eloF* gene is not present, despite some expression showing up in the qPCR. Parts of the sequence show similarity to the other species in this study, but a large portion of the gene

including part of the coding region, is missing. The RNA would be encoding a truncated protein sequence. This suggests *eloF* is not playing an important role in *D. erecta* CHC production, even if it is being expressed. *D. erecta* females do differ from *D. melanogaster* and *D. sechellia* females in the lengths of their predominant CHCs, which typically have longer carbon chains (Cobb and Jallon 1990). The lack of *eloF* expression in *D.erecta* is consistent with the lack of carbon chains of 27 carbons in length. The additional carbons in *D. erecta* suggest another elongase is at work to extend the chains in that species, and that *eloF* is acting only in species whose females produce carbon chains of 27 carbons in length.

Desaturases

Of all desaturases tested in this study, *desatF* is the only gene to also fit my predicted pattern, although there is a minimal amount of expression in *D. sechellia* males. While some dienes are present on the *D. sechellia* male cuticle, dienes account for only 2% of the total CHCs produced by the male (Jallon and David 1987). *Desat1*, although expressed in both sexes of each species, had an interesting pattern of more female expression in monomorphic species and more male expression in dimorphic species. While this pattern of expression does appear sex-biased at first glance, the only statistically significant difference was in *D. melanogaster*. *Desat1* inserts a double bond at the 7th carbon in saturated fatty acids in *D. melanogaster* males and females, resulting in ω7-monoenic fatty acids (Wicker-Thomas et al. 1997). The known role of *desat1* in both males and females is consistent with the expression pattern I observed in this study. A sex-biased pattern would not be expected, given that some CHCs in both males and females contain a double bond at the 7th carbon. *Desat1* also has pleiotropic effects, being involved in fatty acid and lipid metabolism (Ueyama et al. 2005), as well as pheromone production and detection (Bousquet et al. 2012).

My finding of *desat2* expression in both sexes of *D. sechellia* and *D. melanogaster* differs from previous findings that show *desat2* expression to be specific to certain African *Drosophila* populations, not used in this study, producing 5,9-HD (Dallerac et al. 2000). *Desat2* inserts the double bond at the 5th carbon in the geographically specific variant of 7,11-HD. This result could relate to the pleiotropic role of *desat2* in desiccation resistance (Greenberg et al. 2003). The idea of the involvement of *desat1* and *desat2* in inserting two double bonds in dienes has been suggested (Marcillac et al. 2005), however a test of the functional expression of the genes in yeast detected no diene production (Dallerac et al. 2000). Although *desat1* and *desat2* have a role in inserting a double bond at either the 5th or 7th carbon in *Drosophila* fatty acid chains, including pheromonal CHCs, it is unlikely they are involved in differences between *D. simulans* and *D. sechellia*.

EloF gene cluster on 3rd chromosome – CG9458, CG9459, eloF, CG16904, CG8534

Of the other genes in the *eloF* cluster besides *eloF*, none had a clear sex-biased expression pattern. Little is known about the functions of these genes, other than that they are within the elongase family of proteins. Although predicted to have fatty acid elongation activity (Attrill et al. 2016), *CG9458* is unlikely to be involved in sex-specific CHC production because it does not have sex-specific expression. *CG9459*, whose expression was not detected at all in the abdominal tissue, is unlikely to be involved in any pheromonal CHC production. *CG8534* has predicted elongase activity (Attrill et al. 2016), but there is not a clear distinction between male and female activity in the abdomen tissue. *CG16904* is the gene in the *eloF* cluster that has the most potential for a sex-biased role based on its expression pattern. Male bias in *CG16904* expression, could be involved in sex-specific CHC production, if the *CG16904* protein plays a repressive role or acts in some way in males to differentiate the male and female CHC

production pathways. Further functional testing of *CG16904* is needed to determine if this gene is indeed playing a role in sex-specific CHC production.

bond gene cluster on 3rd chromosome – CG5326, CG5278, bond, CG33110, CG6660, CG31141

The slight trend of *CG5326* toward male specificity in dimorphic species combined with the female specificity in *D. mauritiana* make CG5326 a potentially interesting gene to examine. With its predicted elongase activity (Attrill et al. 2016), CG5326 could be involved in male CHC production, and further functional testing could determine if this expression profile is significant for sex-specific CHC production. The same could be said for CG33110, which appears to have a trend favoring male expression. Again, because little is known about the roles of these genes, further functional testing is necessary.

CG6660, CG5278, and CG31141 may be involved in *D. erecta* CHC production. The higher expression of *CG6660* in *D. erecta* females, as opposed to equal sex expression in *D. sechellia* and *D. melanogaster*, is of interest because the primary CHCs of *D. erecta* females are longer in length than those of *D. sechellia* and *D. melanogaster* (Jallon and David 1987). Perhaps *CG6660* is playing a role in elongating *D. erecta* CHCs. Similarly, the higher expression in of *CG5278* and *CG31141* in *D. erecta* males than females, while lacking expression in the other species, signifies some possible role of those genes in the elongation of *D. erecta* male CHCs. Although *D. erecta* male CHCs do not differ greatly from male CHCs of the other species, the prevalence of possible elongase duplications (Howard and Blomquist 2005) makes it is possible that different elongases with similar functions could be used in different species.

Bond has recently been shown to be essential for *Drosophila* male pheromone biosynthesis and fertility (Ng et al. 2015). Silencing *bond* activity severely suppresses male fertility, and fertility can be restored by the ectopic expression of bond in the male reproductive

system (Ng et al. 2015). The known role of *bond* in male pheromone biosynthesis could be consistent with the high male expression in *D. melanogaster* male abdomens, and the low expression in females. Despite *bond's* function in males, it is not likely causing the difference in CHCs between *D. simulans* and *D. sechellia* because the sex-specific expression is not present.

Allele-specific expression of desatF and eloF

Both *cis*- and *trans*- regulatory changes are important in phenotypic evolution, but *cis*-regulatory changes are thought to have a larger influence on morphological changes and species divergence (Wagner and Lynch 2010; Wittkopp and Kalay 2012; Meiklejohn et al. 2014). The use of species hybrids is effective in determining whether variable gene expression is due to *cis*- or *trans*- differences (Wittkopp and Kalay 2012). The differential expression of two alleles in the same cellular conditions in a hybrid points to a *cis*-regulatory effect (Wittkopp and Kalay 2012). My hybrid assay results suggest that a *cis*-regulatory element, such as a promoter or transcription factor binding site, has been altered or lost in *D. simulans* resulting in the failure of expression of the *D. simulans* allele even in the presence of *D. sechellia* trans-activating proteins. The hybrids in this study had a haploid complement of the *D. simulans* genome and of the *D. sechellia* genome. If the cause for differential expression was the presence of an *trans*-activator present in *D. sechellia,* and the loss of that activator in *D. simulans*, then the *D. simulans* allele in the hybrids should have been activated by the *D. sechellia* activator. Because the *D. sechellia* genome was unable to activate expression of the *D. simulans* alleles, this suggests that at least part of the *D. simulans* promoter region has been lost. Another possibility is that mutation has caused a female-specific transcription factor binding site to be gained in *D. sechellia*, resulting in activation of the *D. sechellia* allele, but no activation of the *D. simulans* allele. This gain of a female-specific binding site has occurred for *desatF* in *D. melanogaster* (Shirangi et al. 2009).

Alternatively, a site of repression could have been gained in *D. simulans*, resulting in the lack of expression.

Because the *desatF* and *eloF* coding sequences exist in *D. simulans*, and are very similar to *D. melanogaster* and *D. sechellia*, the dimorphic state may be the ancestral state, with the genes being silenced in *D. simulans*. This scenario requires only one change in expression. If the ancestral state was monomorphic expression, then changes in gene expression would have had to occur in both *D. melanogaster* and *D. sechellia*. In the case of *desatF*, there is relative conservation of codons in the coding sequence among *D. simulans, D. sechellia*, and *D. melanogaster* (Legendre et al. 2008). The promoter regions of *D. simulans desatF*, however, seem to have diverged rapidly, being 10-15% longer than the promoter regions in *D. melanogaster* (Legendre et al. 2008). Several transcription factor binding sites have been found in the 5'-flanking sequence of *desatF*, but further functional testing is needed to determine the roles of these regulatory elements in *desatF* expression (Legendre et al. 2008). One transcription factor binding site in particular, however, the DSX-F site, directly affects transitions from sexual monomorphism to dimorphism (Shirangi et al. 2009), and there have been at least five losses of sexual dimorphism in *D. melanogaster* group (Shirangi et al. 2009). Changes in *cis*-regulatory sequences seem to be the best explanation for how *desatF* expression has rapidly changed from dimorphism to monomorphism.

In addition to the *cis*-regulatory effect on expression differences between *D. simulans* and *D. sechellia* alleles, a *trans*-regulatory effect could explain the reduction in hybrid expression of the *D. sechellia* alleles of both *desatF* and *eloF*. The expression of both genes was greatly reduced in the hybrids relative to *D. sechellia* females, which suggests that two copies of the *D. sechellia* chromosome are necessary to activate full expression of either gene. There could likely

be an epistatic effect of an activator such as a transcription factor, wherein one copy will activate a small amount of expression of the *D. sechellia* allele, but two copies will have a non-additive effect, activating expression of much more than twice the amount of one copy. Thus, while *cis*-regulation could explain expression differences between *D. simulans* and *D. sechellia*, *trans*-regulation could explain expression differences between hybrids and *D. sechellia*.

Conclusion

My results make *desatF* and *eloF* strong candidates for involvement in sex-specific CHC production differences between *D. simulans* and *D. sechellia*, particularly due to the known involvement of both genes in sex-specific CHC production in the closely related species *D. melanogaster* (Legendre et al. 2008). If *desatF* and *eloF* are acting in the *D. sechellia* pathway to the female CHC 7,11-HD, then they are examples for two specific genes directly involved in premating reproductive isolation. This sex-specific expression difference between *D. simulans* and *D. sechellia* is a prime example of how changes in the expression of genes in the biosynthesis pathway of a reproductive trait can cause or influence reproductive isolation between species. While it is unknown if this expression difference originated before or after *D. simulans* and *D. sechellia* diverged, it is clear that this sex-specific expression of *desatF* and *eloF* is playing a role in maintaining species boundaries, and could have influenced speciation. To test the functional effects of *desatF* and *eloF* on mating behavior, further experimentation on courtship is needed, which is the focus of the third chapter of this study.

Literature Cited

Attrill, H., K. Falls, J. L. Goodman, G. H. Millburn, G. Antonazzo, A. J. Rey, S. J. Marygold, and C. FlyBase. 2016. FlyBase: establishing a Gene Group resource for *Drosophila melanogaster*. Nucleic Acids Research 44:D786-792.

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

Bousquet, F., T. Nojima, B. Houot, I. Chauvel, S. Chaudy, S. Dupas, D. Yamamoto, and J. F. Ferveur. 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 of the United States of America 109:249-254.

Chertemps, T., L. Duportets, C. Labeur, R. Ueda, K. Takahashi, K. Saigo, and C. Wicker-Thomas. 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 J. M. Jallon. 1990. Pheromones, mate recognition and courtship stimulation in the *Drosophila melanogaster* species subgroup. Animal Behaviour 39:1058-1067.

Dallerac, R., C. Labeur, J. M. Jallon, D. C. Knippie, W. L. Roelofs, and C. Wicker-Thomas. 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 of the United States of America 97:9449-9454.

Fang, S., C. T. Ting, C. R. Lee, K. H. Chu, C. C. Wang, and S. C. Tsaur. 2009. Molecular evolution and functional diversification of fatty acid desaturases after recurrent gene duplication in *Drosophila*. Molecular Biology and Evolution 26:1447-1456.

Ferveur, J. F. 1997. The pheromonal role of cuticular hydrocarbons in *Drosophila melanogaster*. Bioessays 19:353-358.

Garrigan, D., S. B. Kingan, A. J. Geneva, P. Andolfatto, A. G. Clark, K. R. Thornton, and D. C. Presgraves. 2012. Genome sequencing reveals complex speciation in the *Drosophila simulans* clade. Genome Research 22:1499-1511.

Gleason, J. M., J. M. Jallon, J. D. Rouault, and M. G. Ritchie. 2005. Quantitative trait loci for cuticular hydrocarbons associated with sexual isolation between *Drosophila simulans* and *D. sechellia*. Genetics 171:1789-1798.

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

Gleason, J. M. and M. G. Ritchie. 2004. Do quantitative trait loci (QTL) for a courtship song difference between *Drosophila simulans* and *D. sechellia* coincide with candidate genes and intraspecific QTL? Genetics 166:1303-1311.

Graze, R. M., O. Barmina, D. Tufts, E. Naderi, K. L. Harmon, M. Persianinova, and S. V. Nuzhdin. 2007. New candidate genes for sex-comb divergence between *Drosophila mauritiana* and *Drosophila simulans*. Genetics 176:2561-2576.

Greenberg, A. J., J. R. Moran, J. A. Coyne, and C. I. Wu. 2003. Ecological adaptation during incipient speciation revealed by precise gene replacement. Science 302:1754-1757.

Hackett, J. L. 2011. Genetic contributions of *desatF* and *eloF* to courtship mating behavior and cuticular hydrocarbon production in *Drosophila simulans* and *D. sechellia*. Molecular Biosciences. Master of Arts Thesis, University of Kansas, Lawrence, KS.

Hellemans, J., G. Mortier, A. De Paepe, F. Speleman, and J. Vandesompele. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biology 8:R19.

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

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

Jallon, J. M. and J. R. David. 1987. Variations in cuticular hydrocarbons among the eight species of the *Drosophila Melanogaster* subgroup. Evolution 41:294-302.

Kastanis, P., E. Eliopoulos, G. N. Goulielmos, S. Tsakas, and M. Loukas. 2003. Macroevolutionary relationships of species of *Drosophila melanogaster* group based on mtDNA sequences. Molecular Phylogenetics and Evolution 28:518-528.

Legendre, A., X. X. Miao, J. L. Da Lage, and C. Wicker-Thomas. 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.

Mackay, T. F. 2004. The genetic architecture of quantitative traits: lessons from *Drosophila*. Current Opinion in Genetics & Development 14:253-257.

Marcillac, F., Y. Grosjean, and J. F. Ferveur. 2005. A single mutation alters production and discrimination of *Drosophila* sex pheromones. Proceedings of the Royal Society B: Biological Sciences 272:303-309.

Meiklejohn, C. D., J. D. Coolon, D. L. Hartl, and P. J. Wittkopp. 2014. The roles of *cis*- and *trans*-regulation in the evolution of regulatory incompatibilities and sexually dimorphic gene expression. Genome Research 24:84-95.

Ng, W. C., J. S. Chin, K. J. Tan, and J. Y. Yew. 2015. The fatty acid elongase *Bond* is essential for *Drosophila* sex pheromone synthesis and male fertility. Nature Communications 6:8263.

Peirson, S. N., J. N. Butler, and R. G. Foster. 2003. Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. Nucleic Acids Research 31:e73.

Selifonova, O., F. Valle, and V. Schellenberger. 2001. Rapid evolution of novel traits in microorganisms. Applied and Environmental Microbiology 67:3645-3649.

Shirangi, T. R., H. D. Dufour, T. M. Williams, and S. B. Carroll. 2009. Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*. PLoS Biology 7 (8): e1000168.

Smadja, C. and R. K. Butlin. 2009. On the scent of speciation: the chemosensory system and its role in premating isolation. Heredity (Edinb) 102:77-97.

Ueyama, M., T. Chertemps, C. Labeur, and C. Wicker-Thomas. 2005. Mutations in the desat1 gene reduces the production of courtship stimulatory pheromones through a marked effect on fatty acids in *Drosophila melanogaster*. Insect Biochemistry and Molecular Biology 35:911-920.

Wagner, G. P. and V. J. Lynch. 2010. Evolutionary novelties. Current Biology 20:R48-52.

Wicker-Thomas, C., C. Henriet, and R. Dallerac. 1997. Partial characterization of a fatty acid desaturase gene in *Drosophila melanogaster*. Insect Biochemistry and Molecular Biology 27:963-972.

Williams, T. M., J. E. Selegue, T. Werner, N. Gompel, A. Kopp, and S. B. Carroll. 2008. The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. Cell 134:610-623.

Wittkopp, P. J., B. K. Haerum, and A. G. Clark. 2004. Evolutionary changes in *cis* and *trans* gene regulation. Nature 430:85-88.

Wittkopp, P. J. and G. Kalay. 2012. *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nature Reviews Genetics 13:59-69.

Zauner, H., G. Begemann, M. Mari-Beffa, and A. Meyer. 2003. Differential regulation of *msx* genes in the development of the gonopodium, an intromittent organ, and of the "sword," a sexually selected trait of swordtail fishes (Xiphophorus). Evolution & Development 5:466-477.

Table 2.1 QPCR Primer Sequences

Gene ID, name	Forward Primer (5' → 3')	Reverse Primer (5' → 3')
CG7923, *desatF*	CCTGAACACTTTGGCCTTCC	ATTTGCTTGCCCTTCTCCAC
CG5887, *desat1*	ACATCATCGCCTTCGGTTAC	CGTGTTGAAGATGACCAGNA
CG5925, *desat2*	GTGACCTGACCACCGACAG	GCCATATAGAGCAGCCAGGT
CG9458	TGGCCACCTATCTGNTCTTC	GGGTCCCAGCATAAAGTGNA
CG9459	GCAGANNCGAAAGCCNTACA	CAGGANCCAGCATGAAGTG
CG16905, *eloF*	CAACATATTCCAGATCCTTTACAA	ATCCTTATATTTGTGATCCATCG
CG8534	GTCCGCCTGCCACTCATC	CTGCATAATGTTGTATGCCCTGA
CG5326	CCGTGTGTGCCTTCATTG	TCTGCAGGATGGTGATGTACT
CG5278	TCAACGCAACACAGGTNGAC	TGCGATCCTTCATAAACTTGG
CG33110	GCTGCACGTGTACCATCACT	CCACAGGAACTTGGCGTACT
CG6660	TGCCACGATATTCGTCATTG	TCTTCTTGCGCAGCACTATG
CG31141	AGGAAGTTNATGGAGCATCG	AGCGAAAGTTGTACGGTTGG
CG6921, *bond*	GAAGATCGGACCCGAGTACA	GATCGACGACATCACATTGC
CG7939, *rp49*	ATGCTAAGCTGTCGCACAAA	ACGTTGTGCACCAGGAACTT
CG15573, *Femcoat*	GGACCACAACTAATGCTGCTG	TCTTTCTTTTCGTCCCACCA
CG8982, A*cp26Aa*	GAACCTGATTTTGTTATGCTCTCA	TGGGAAGGAAGAGTGGAAGA

qPCR reaction mix: 10 µl SYBR® Select Master Mix (Life Technologies), 0.1 µl (150 nM) forward primer, 0.1 µl (150 nM) reverse primer, 7.3 µl dH_20, 2.5 µl cDNA (50 ng).

qPCR reaction conditions: Initial stage of 50˚C for 2 minutes, denaturation at 95˚C for 2 minutes, then 40 qPCR cycles of 95˚C for 15 seconds, annealing/extension at 60˚C for 1 minute.

Table 2.2. TAQMAN Probe Sequences and Dyes.

Probe specific species/gene	Probe sequence	Sequence location	5' Probe dye
D. simulans eloF	TGAAAGCCTACCAAATCAGCTG**C**ATTGTCAG	2nd exon	6FAM
D. sechellia eloF	TGCTGAAAGCCTACCAAATCAGCTGTATTGTC	2nd exon	VIC
D. simulans desatF	TGGCTGGTTCTTTGCCCACATTGGAT	exon	6FAM
D. sechellia desatF	TGGCTGGTTCTTTGCCCACATAGGAT	exon	VIC

Probe sequences are given 5' to 3', with nucleotide differences underlined for each gene.

qPCR conditions: Initial stage of 50˚C for 2 minutes, denaturation at 95˚C for 10 minutes, then 40 PCR cycles of 95˚C for 15 seconds, annealing/extension at 58˚C (*eloF*) or 60˚C (*desatF*) for 15 seconds, and a second extension step at 74˚C for 30 seconds.

Figure 2.1

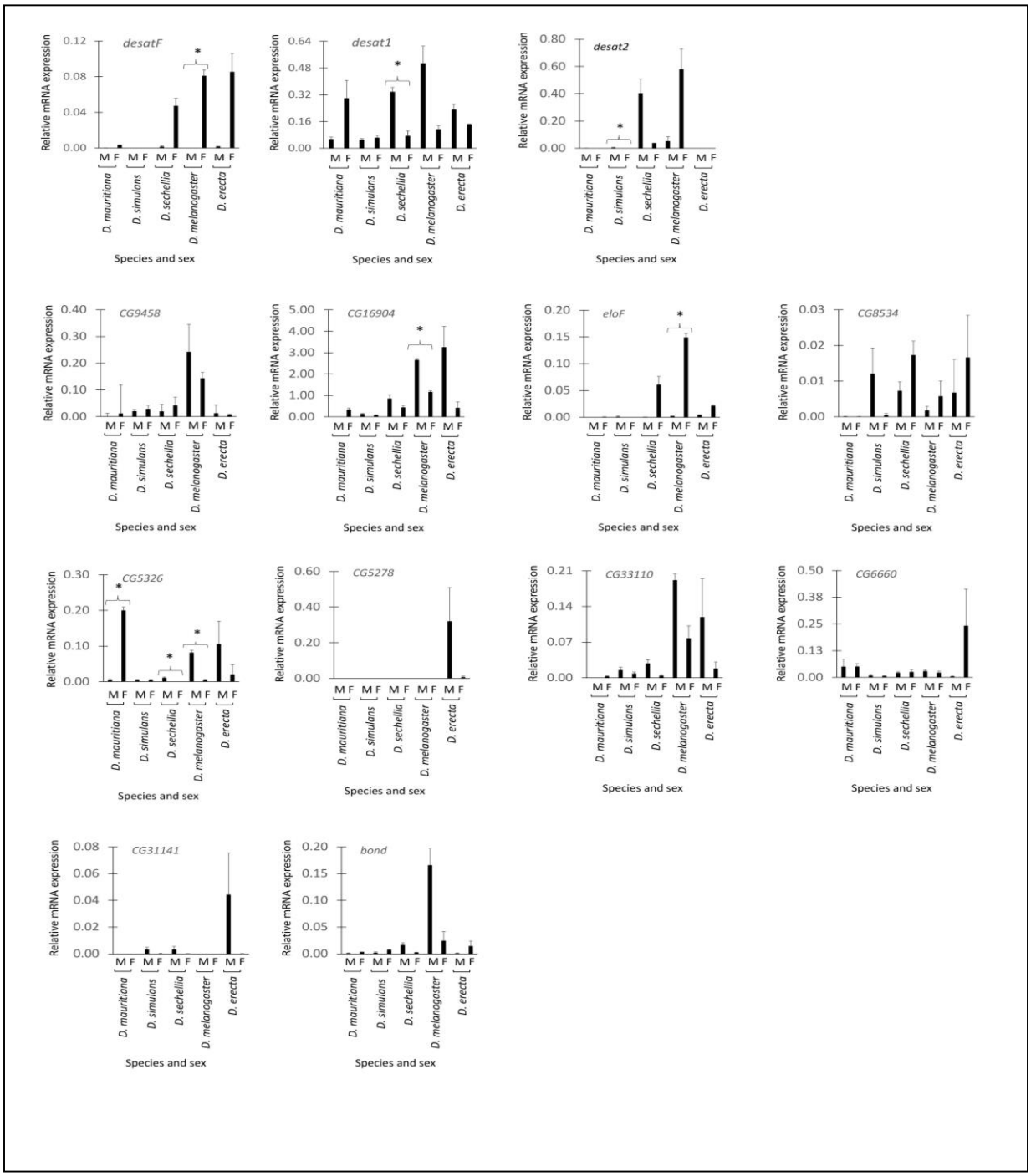
desatF
desat1
desat2
CG9458
CG16904
eloF
CG8534
CG5326
CG5278
CG33110
CG6660
CG31141
bond
Relative mRNA expression
Species and sex
M F
D. mauritiana
D. simulans
D. sechellia
D. melanogaster
D. erecta

Figure 2.1 Quantitative RT-PCR of candidate genes among each sex of *D. simulans*, *D. mauritiana*, *D. sechellia*, *D. melanogaster*, and *D. erecta* shows dimorphic female expression only in *desatF* and *eloF*. To determine whether candidate genes are likely involved in CHC production differences between *D. simulans* and *D. sechellia*, I performed qPCR to test the prediction that genes involved in this CHC difference are only be expressed in dimorphic females. The Y axis of each graph shows expression amount relative to the ribosomal protein rp49. The X axis of each graph shows expression in each species, separated by sex within species. Student's t-tests were performed for differences between sexes within species (*=t-test $P<0.05$).

Figure 2.2

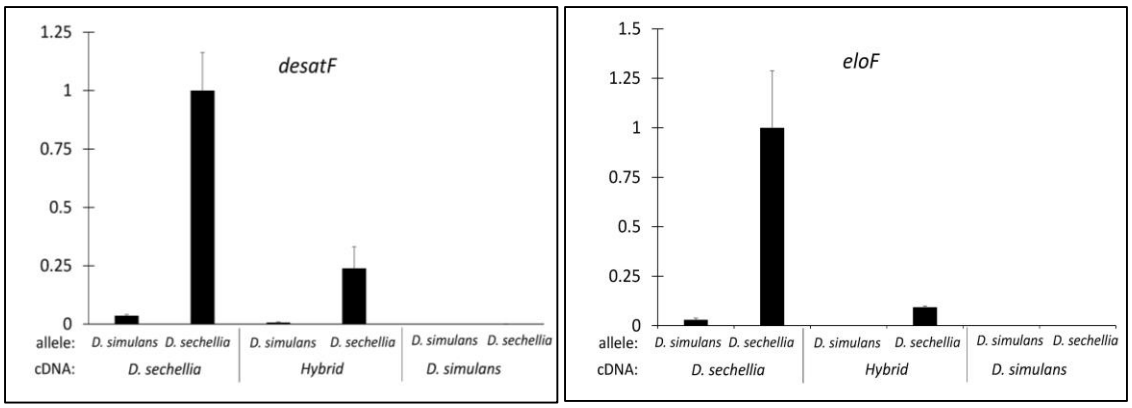
desatF
1.25
1
0.75
0.5
0.25
0
allele:
D. simulans
D. sechellia
D. simulans
D. sechellia
D. simulans
D. sechellia
cDNA:
D. sechellia
Hybrid
D. simulans
eloF
1.5
1.25
1
0.75
0.5
0.25
0
allele:
D. simulans
D. sechellia
D. simulans
D. sechellia
D. simulans
D. sechellia
cDNA:
D. sechellia
Hybrid
D. simulans

Figure 2.2 Allele-specific qPCR of *desatF* and *eloF* in *D. simulans/D. sechellia* hybrid females shows lack of expression of the *D. simulans* allele and much less expression of the *D. sechellia* allele compared to *D. sechellia* wild type females.

For *desatF* and *eloF*, I designed Taqman probes specific for either the *D. simulans* or *D. sechellia* allele. Primers were designed for each gene to amplify both species. Samples included *D. simulans*, *D. sechellia*, and hybrid cDNA. For comparison, the expression amount of *desatF* or *eloF* in *D. sechellia* females was set to 1.0, and all other samples were compared as a fold difference to *D. sechellia* females. The Y axis shows relative fold difference.

Chapter 3

Introgression of *desatF* and *eloF* Alleles Affects Gene Expression and Courtship between *D. simulans* and *D. sechellia*.

Introduction

Pre-mating reproductive isolation can maintain biological mating boundaries between species (Dobzhansky 1935). One essential component of pre-mating reproductive isolation is the ability of individuals in a population to recognize the species of potential mating partners (reviewed in Greenspan and Ferveur 2000). In *Drosophila*, courtship behaviors facilitate mate recognition, and can lead to the isolation of populations from one another (reviewed in Greenspan and Ferveur 2000). Thus, the elucidation of genes and gene expression affecting courtship behavior can provide insight how genes influence reproductive isolation.

Cuticular hydrocarbons (CHCs) strongly influence mate choice in Drosophila

CHCs strongly influence mate. Applying 7,11-HD to *D. simulans* females prevents courtship from *D. simulans* males (Coyne et al. 1994). Preventing 7,11-HD production by the ablation of oenocytes in *D. melanogaster* elicits courtship from *D. simulans* males (Billeter et al. 2009). Treating ablated females artificially with 7,11-HD restores the courtship barrier with *D. simulans* males. Removing CHCs and perfuming with artificial CHCs breaks and restores courtship barriers in other *Drosophila* species as well (Mas and Jallon 2005). Genetically feminizing CHCs in *D. melanogaster* males causes courtship to occur from other males (Ferveur et al. 1997; Wang et al. 2011). These experiments demonstrate that CHCs play a vital role in establishing species and sex identity in *Drosophila*.

The use of near-isogenic lines in studying quantitative traits

Behavior is typically a quantitative trait. Variation in quantitative traits among species suggests differences at multiple quantitative trait loci, each with small effects on the overall

phenotype (Mackay 2001). Variation of quantitative traits is particularly important between sexes, as complex behavior often differs between males and females (Anholt and Mackay 2001).

One effective method for studying the effects of variation at particular QTL is the use of genomic introgression to make near-isogenic lines (NILs; Anholt and Mackay 2001). NILs primarily consist of one genome, but contain a small, introgressed genomic region from a second genome at a particular, controlled spot. NILs have been used to study quantitative traits in both animals (Lyman and Mackay 1998; Ashton et al. 2001; Robin et al. 2002) and plants (Keurentjes et al. 2007). The effectiveness of NILs is in their power to detect differential allelic effects at a particular QTL. That power can be increased by making multiple NILs all with the same introgression (Keurentjes et al. 2007). Furthermore, by using two NILs with different introgressions to make a NIL containing both introgressions, epistatic interactions can be detected. The construction of NILs is particularly useful in observing small-effect differences of the introgressed region (Anholt and Mackay 2001). Phenotypic variation between NILs and wildtype species can confirm that the introgressed region is affecting the trait, although it does not identify specific genes within the region having an effect (Anholt and Mackay 2001).

Chapter Goals

Here I test the functional roles of *desatF* and *eloF* expression on courtship between *D. simulans* and *D. sechellia* by introgressing *D. simulans* alleles of each gene region into a *D. sechellia* genomic background. As demonstrated in Chapter 2, *desatF* and *eloF* have a pattern of expression consistent with involvement in female-specific CHC production. I introgressed *D. simulans* alleles of *desatF* and *eloF* into a *D. sechellia* genetic background through the production of near isogenic lines (NILs). I observed the presence or absence of expression of *desatF* and *eloF* in the NILs, and performed courtship experiments to determine the effects of the

introgressions. For the gene expression assay, females of lines homozygous for the *D. simulans* allele of *desatF* (termed "*desatF* NILs") were predicted to lack expression of *desatF*, and females of lines containing the *D. simulans* allele of *eloF* (termed "*eloF* NILs") were predicted to lack expression of *eloF*. Females of a line containing the *D. simulans* allele of both *desatF* and *eloF* (termed "double NIL") were predicted to lack expression of both *desatF* and *eloF*. For the courtship experiments, NIL females were predicted to be courted by *D. simulans* males more than *D. sechellia* females, but less than *D. simulans* females. Such results demonstrate that eliminating *desatF* and/or *eloF* expression in *D. sechellia* females can alter the courtship behavior normally elicited from *D. simulans* males, and suggest that these genes are involved in maintaining a reproductive boundary between the two species.

Materials and Methods

Fly stocks and culturing

Cultures of *D. sechellia* strain David 4A (described in Gleason and Ritchie 2004), *D. simulans* strain f^2*;nt,pm,e (*described in Gleason and Ritchie 2004*)* and *D. simulans* strain 14021-0251.169 (San Diego Stock Center, from now on referred to as sim169) were maintained on standard cornmeal-molasses food in 25 x 95 mm polystyrene vials in populations of approximately twenty flies. The flies were kept at 25°C on a 12 hour light/12 hour dark cycle.

DNA extraction and genotyping

Whole body genotyping was performed using the single fly prep method (Gloor 1992). Wing genotyping was also performed using a modification of the whole body protocol (Gleason et al. 2004). Forceps were used to remove both wings of an individual as close to the body as

possible. While the wings were being genotyped, flies were housed individually in a 16.5 x 95 mm polystyrene vial with about 2 mL of standard cornmeal-molasses media. Following PCR using primers and conditions given in Table 3.1, agarose gel electrophoresis of the PCR products was performed to determine the genotype of the *desatF* and *eloF* alleles, which gave different sized bands to differentiate genotypes.

Construction of near-isogenic lines

I used marker-assisted selection to produce near-isogenic lines homozygous for the *D. simulans* (strain f^2*;nt,pm,e*) alleles of either *desatF* or *eloF* in an otherwise mostly *D. sechellia* genomic background. The two genes were treated separately to introgress a minimum amount of *D. simulans* around each gene into *D. sechellia*. I first crossed ten virgin *D. simulans* females with ten virgin *D. sechellia* males. F1 females were backcrossed to *D. sechellia* males. F2 females were mated to *D. sechellia* males, then removed for whole body DNA extraction and genotyping for *desatF* and *eloF*. Vials containing eggs from females that were heterozygous at either gene were kept and the female progeny were used in the subsequent cross to *D. sechellia* males.

For all subsequent generations, females were mated with three *D. sechellia* males before genotyping and reserving progeny of females that were heterozygous for either gene. The process was continued for 21-28 generations to minimize the size of the introgressed region. Because progeny production was low, some generations were established by backcrossing a male to three *D. sechellia* females.

Homozygous lines were established by crossing heterozygous siblings. Heterozygous parents were chosen by wing genotyping. The resulting progeny were wing genotyped and homozygous individuals were crossed to establish a line. From this procedure, I established five

lines homozygous for the *D. simulans eloF* and one line homozygous for the *D. simulans desatF*. The *eloF* lines were split at generation three and thus should represent independent introgression intervals.

To establish a line homozygous for the *D. simulans* allele of both *desatF* and *eloF*, I crossed ten virgin females from *eloF* line 1 with ten virgin males from the *desatF* line. F1 siblings from this cross were then mated, and F2 individuals were wing genotyped. Because recombination does not occur in males, and the two genes are both on the 3rd chromosome, none of the F2 progeny was homozygous for both genes. F2 females homozygous for *D. simulans* at one gene (either *desatF* or *eloF*) and heterozygous at the other gene were then mated to F2 males of the same genotype. F3 siblings that were homozygous *D. simulans* at both genes were then mated to each other. This cross resulted in F4 progeny that were all homozygous *D. simulans* at both *desatF* and *eloF*. These progeny established a double homozygous line.

Genotyping markers adjacent to desatF and eloF

To determine the size of the introgressions, I genotyped the region around each gene. I extracted DNA from five individuals of each NIL and genotyped them using primers placed in adjacent genes. This produced PCR products of different allele lengths between *D. simulans* and *D. sechellia* (Table 3.1) that were visualized on an agarose gel for determine the species origin of each locus.

Measurement of desatF and eloF expression

Because *desatF* and *eloF* are not expressed in *D. simulans* (Chapter 2), I checked for expression from the introgressed genes. RNA was extracted and cDNA synthesized from females using the method described in Chapter 2. RT-PCR was performed on each cDNA sample using

primers designed for each gene (Table 3.2). The ubiquitously expressed ribosomal protein gene *rp49* (Krupp et al. 2008) was used as a positive control. Following PCR, the samples were electrophoresed and visualized with ethidium bromide on a 2% agarose gel, with genomic DNA samples added as a size marker. Presence of expression was scored when a band was observed of the appropriate size for cDNA. Absence was scored if the band was not present, yet *rp49* was amplified. Two RNA/cDNA preparations were made for each species and sex as biological replicates, and two PCR reactions were performed on each sample as technical replicates.

Courtship experiments

To test the effects of the introgressions on *D. simulans* male courtship, courtship experiments were performed with each NIL, *D. sechellia* line David 4A, and *D. simulans* line sim169. To standardize the cultures, five virgin females and five virgin males were mated in a 25 x 95 mm polystyrene vial. F1 virgin females and males were housed individually in a 16.5 x 95 mm polystyrene vial with standard cornmeal-molasses food for seven days before use in the trials. All trials were performed at 25˚C within four hours post lights on.

To determine if the alterations to *desatF* and *eloF* gene regions in the NILs elicit courtship, I first carried out the trials with *D. simulans* males. For each pairing observed, one male and one female were aspirated into circular clear chambers (25 mm diameter, 12 mm height). Two chambers were video recorded with a single video microscope (Veho VMS-001) and two microscopes recorded chambers simultaneously for a total of four chambers. The chambers were set up in the following manner: 1. NIL female strain 1 with *D. simulans* male, 2. NIL female strain 2 with *D. simulans* male, 3. *D. simulans* female with *D. simulans* male, 4. *D. sechellia* female with *D. simulans* male. Positioning of the chambers under the two cameras was randomized for each trial using a random number generator. All females were aspirated first into

the chambers. Once each female was in the chamber, video recording began. The males were then added to each chamber. Pairs were recorded for 1 hour using Windows Movie Maker (Microsoft). After the fly identifications for each chamber were recorded, the video was given a 7-digit random number using a random number generator.

When a set of trials was completed, analysis of video was performed blindly with respect to the identity of the flies in each mating chamber. The beginning of the observation period was marked by the addition of the male. The time when courtship was initiated was scored when the male approached the female, oriented and tapped the female abdomen. If courtship was not initiated within thirty minutes, the trial was ended. Once courtship was initiated, the cumulative time the male courted was measured by including all times the male engaged in tapping, singing, licking and mounting the female (reviewed in Greenspan and Ferveur 2000). Observations continued for thirty minutes or until copulation began, whichever occurred first.

Several parameters were calculated for each pair. Courtship latency (CL) was the time between the addition of the male and the time the male first initiated courtship. If the male did not court, then CL was recorded as 1801 seconds (one second longer than the observation period). The courtship index (CI) was calculated as the total time the male spent courting from the initiation of courtship to the end of the trial divided by the total time from the initiation of courtship to the end of the trial, and thus varies between 0 (no courtship after initiation) to 1 (continuous courtship).

To test whether the *desatF* and *eloF* introgressions decrease courtship of the NILs from *D. sechellia* males relative to *D. sechellia* females, I performed courtship trials with *D. sechellia* males. The procedure for these trials was the same described above, using *D. sechellia* males rather than *D. simulans* males.

Statistical analyses of courtship trials

At least ten trials were performed for each NIL with controls. Each NIL had its own set of control pairings (with *D. simulans* and *D. sechellia* females), and statistical tests were performed for each NIL separately. Courtship index was not calculated when males failed to initiate courtship. A two-tailed Student's t-test was performed to test for significant differences between sample groups (NIL, *D. simulans, D. sechellia* females) in courtship latency and courtship index when paired with either a *D. simulans* or *D. sechellia* male. Copulation occurrences were recorded for each trial, but the number of copulations was so low that statistical analysis of copulation differences was not performed. For overall courtship calculations, females of NILs, *D. simulans*, and *D. sechellia* were combined by species and averaged. It should be noted that the overall data was obtained by combining data from separate groups of trials.

Results

Genetic makeup of Near-Isogenic Lines

I made five distinct lines homozygous for the *D. simulans* allele of *eloF* in a mostly *D. sechellia* genomic background (hereafter designated "eloF1-eloF5"), one line homozygous for the *D. simulans* allele of *desatF* in a mostly *D. sechellia* genomic background (designated "desatF"), and one line homozygous for both the *D. simulans eloF* and the *D. simulans desatF* alleles in a mostly *D. sechellia* genomic background (hereafter, "double line"). I found that each of the five *eloF* NILs had similar introgressions in a 44 kb region around *eloF* on chromosome 3R (Figure 3.1). For each of the five *eloF* lines, the markers *CG16904*, and *CG9459* to the left of *eloF* were from *D. simulans*, and *Teh1* was *D. sechellia* indicating that the recombination

breakpoint was at most 14 kb away. To the right of *eloF*, *CG8534* and *CG8516* were from *D. simulans*, and *MtnA* was *D. sechellia*, indicating that the recombination breakpoint was at most 30 kb away. Each *eloF* NIL was also genotyped at *desatF* and carried the *D. sechellia* allele.

For *desatF* (Figure 3.1), the *Tna* allele was from *D. sechellia*, indicating that the recombination breakpoint was at most 153 kb away to the left. To the right, *CG43693* was *D. simulans* and *CG8534* was *D. sechellia*, indicating that the recombination breakpoint was at most 19 kb away. *Tna* was the closest marker I tested to the left of *desatF*, so the introgression size could be much smaller. The desatF NIL was genotyped at *eloF* and found to carry the *D. sechellia* allele.

Expression of desatF and eloF in Near-Isogenic Lines (NILs)

I tested females of each NIL for the expression of *desatF* and *eloF* mRNA. Females of each of the five *eloF* lines did not express *eloF* but did express *desatF* (data not shown). The *desatF* line did not express *desatF* but did express *eloF* (data not shown). The double homozygous line (carrying the *D. simulans* allele of both *desatF* and *eloF*) did not express either *desatF* or *eloF* (data not shown). Thus, the NIL females are different from *D. sechellia* females in the expression of at least *desatF* or *eloF*.

Courtship by D. simulans males

Because *D. simulans* males discriminate between conspecifics and heterospecifics, I compared the courtship of *D. simulans* males toward NIL females with courtship directed at *D. simulans* and *D. sechellia* females. Not all males initiated courtship (data not shown). CL from *D. simulans* males toward each NIL, *D. simulans* and *D. sechellia* females did not differ (Figure 3.2) with the exception of initiation toward *D. sechellia* and *D. simulans* females in the trials

including eloF5, when *D. simulans* females were courted more quickly than *D. sechellia* females. Thus, *D. simulans* males do not discriminate females before initiating courtship.

D. simulans males courted *D. simulans* females vigorously, and displayed almost no courtship activity towards *D. sechellia* females across all trials (Figure 3.3). Courtship displayed towards the NILs was variable but all NILs were courted less than *D. simulans*. Three of the NILs, eloF1, eloF4 and eloF5, were courted significantly more than *D. sechellia* females (t-test P=0.004, 0.01, 0.007 respectively). The eloF2 and eloF3 lines, as well as the desatF line, were courted as much as *D. sechellia* females (t-test P=0.09, 0.06, 0.08 respectively). Thus, *D. simulans* males selectively court females once courtship has begun.

Courtship by D. sechellia males

Because *D. sechellia* males do not discriminate between conspecifics and heterospecifics (Cobb and Jallon 1990), I compared the courtship of *D. sechellia* males toward NIL females with courtship directed at *D. simulans* and *D. sechellia* females. Not all males initiated courtship (data not shown). CL from *D. sechellia* males toward each NIL, *D. simulans* and *D. sechellia* female did not differ (Figure 3.4). Thus, *D. sechellia* males do not initially discriminate among potential females.

D. sechellia males displayed a similar amount of courtship toward *D. simulans* and *D. sechellia* females (Figure 3.5). The group of courtship trials with eloF4 and 5 were the only groups in which *D. sechellia* females were courted significantly more than *D. simulans* females (t-test, P= 0.03, 0.03 respectively). Courtship displayed toward NILs was variable, however, only eloF2 females were courted significantly more than *D. simulans* females by *D. sechellia* males (t-test, P=0.02). eloF1 and 3, as well as the desatF line, had no difference in courtship from either *D. sechellia* or *D. simulans* females, and even the control *D. sechellia* and *D.*

simulans females did not have significantly different CI indices. The group of courtship trials with *eloF*4 was the only group in which *D. sechellia* females were courted more than both the NIL and *D. simulans* females (t-test, P=0.04, 0.03 respectively). Trials with eloF5 had a significant difference among *D. sechellia* and *D. simulans* females (t-test, P=0.03), but the eloF line did not differ from either. Thus, *D. sechellia* males discriminate minimally once courtship has begun and the introgressions have not introduced any genomic regions with a negative effect on courtship.

Overall courtship latency and courtship index

The measurements for overall CL and CI were obtained by combining groups of females from different trial sets. This allowed comparisons over all females, but also hid any block effects that may have been present in different sets. Males of *D. simulans* courted more quickly than *D. sechellia* males overall (Table 3.3, t-test, P<0.001). The mean CL of *D. sechellia* males toward all females combined was nearly twice the CL of *D. simulans* males to all females combined. Thus, *D. simulans* males show a propensity to initiate courtship more quickly than *D. sechellia* males. When paired with conspecifics, *D. simulans* males courted more actively than *D. sechellia* males (Table 3.3, t-test, P<0.001). However, when paired with heterospecific females, *D. sechellia* males courted more than *D. simulans* males, which failed to court *D. sechellia* females (Table 3.3, t-test, P=0.001). Thus, *D. simulans* males are more vigorous courters than *D. sechellia* males and highly prefer females of their own species. *D. sechellia* males and *D. simulans* males courted the NILs equally (Table 3.3, t-test, P=0.49), suggesting that the introgressions did not affect any additional female courtship signals.

Discussion

My results suggest that the differential gene expression of *desatF* and *eloF* between *D. simulans* and *D. sechellia* females is contributing to a reproductive barrier between the species. Removing the expression of either *eloF* or *desatF* expression from primarily *D. sechellia* females elicits some courtship from *D. simulans* males. Thus the premating reproductive isolation that normally exists between the species can be affected by altering the expression of a single gene region. It should be re-stated, however, that the introgression regions did not contain only *desatF* and *eloF*, and other genes within the regions cannot be ruled out in having an effect.

Introgression sizes and implications for the results

Because I did not detect differences in the introgression breakpoints among the eloF NILs, the differences in courtship toward each line are possibly due to undetected differences in breakpoints, other untested genomic differences, or a result of small sample sizes. Although the high number of generations in the construction of the NILs suggests that the *D. simulans* genome outside the introgression region should be completely lost, it is possible that some *D. simulans* regions were retained. This could be investigated further by genotyping markers spread throughout the genome.

Expression of desatF and eloF in near-isogenic lines

My findings that introgressed *eloF* and *desatF* alleles both lacked expression supports the idea that the *D. simulans* alleles of *eloF* and *desatF* are effectively turned off, as I showed in Chapter 2. The expression of the *D. sechellia* alleles of both genes is similar to the expression profiles of *desatF* (Legendre et al. 2008) and *eloF* (Chertemps et al. 2007) in *D. melanogaster*

females, which also produce 7,11-HD. Having a *D. sechellia* genomic background does not activate expression of the *D. simulans* alleles. This lack of *eloF* and *desatF* expression in what are essentially *D. sechellia* females suggests that *cis*-regulatory differences in the *D. simulans* alleles are preventing the expression of the genes. This lends support to the argument that *cis*-acting regulatory changes play a large role in interspecific expression differences and in species divergence (Wittkopp et al. 2004). The next step should be to examine the CHC types produced by the NILs to determine what effects *desatF* and *eloF* expression are having on CHC structure.

Courtship latency

These species are asymmetrically sexually isolated because *D. simulans* males will not court *D. sechellia* females, but *D. sechellia* males will court *D. simulans* females (Cobb and Jallon 1990). I measured the effects on courtship of introgression of the inactive *D. simulans* alleles of each gene into a *D. sechellia* genomic background. Courtship latency was not affected by the introgressions for either species, implying that CHCs affected by the introgressed regions do not play a role in courtship until the male physically engages with the female by tapping her abdomen, which allows gustatory receptors on the male's forelegs to detect any CHCs present. Other courtship cues involving senses such as vision or hearing could give information to the male before he approaches the female, thus potentially increasing CL if the signals are perceived negatively, or decreasing CL if the signals are perceived positively.

Courtship index

The removal of 7,11-HD from *D. melanogaster* females can lessen or prevent courtship behavior from *D. melanogaster* males, and elicit courtship from *D. simulans* males (Billeter et al. 2009; Shirangi et al. 2009). In addition, manipulation of *desatF* and *eloF* expression has been

shown to alter courtship behavior in *D. melanogaster* (Chertemps et al. 2006; Chertemps et al. 2007; Legendre et al. 2008). Elimination of *eloF* expression in *D. melanogaster* females increases CL and decreases copulation attempts from wildtype *D. melanogaster* males (Chertemps et al. 2007). Elimination of *desatF* expression in *D. melanogaster* females leads to courtship from *D. simulans* males, similar to the courtship *D. simulans* males give to *D. simulans* females (Legendre et al. 2008). Here I found that the introgression of the *D. simulans* region around *eloF* in *D. sechellia* females can be enough to elicit some courtship from a *D. simulans* male. All five of the eloF NILs were courted much less than wildtype *D. simulans* females though three were courted significantly more than wildtype *D. sechellia* females. This suggests that removal of the *eloF* introgression region alone is not sufficient to elicit the same amount of courtship given to a wildtype *D. simulans* female, but can be enough to receive more courtship than a wildtype *D. sechellia* female. The cause of increased courtship of some NILs relative to *D. sechellia* is presumably the presence of shorter CHCs, which could be tested for by measuring the CHC profiles of the NIL females. The decreased courtship of the eloF NILs relative to *D. simulans* females could be explained by expression of *desatF*, which could be affecting the CHC structure enough to still warrant some aversion from *D. simulans* males.

My results suggest that while some variation in courtship did occur in the *eloF* NILs, the courtship of *D. simulans* males can be altered by the introgression of the *eloF* region. If *eloF* is affecting 7,11-HD production, then removing *eloF* expression should greatly reduce or even eliminate 7,11-HD production. If 7,11-HD is the primary constraint for *D. simulans* courtship of *D. sechellia*, then elimination of 7,11-HD production could affect this barrier to courtship.

In the one *desatF* NIL, the introgression was not enough to cause more courtship of the NIL compared to the *D. sechellia* females, though some courtship did occur. However, I was

only able to make one *desatF* line, which could be limiting my ability to detect significant differences in courtship toward *desatF* NIL and *D. sechellia* females, if they exist. Not all of the eloF lines had a significant difference from *D. sechellia*, so it is possible the sample size for *desatF* simply was not large enough. Further investigation of courtship after the removal of *desatF* expression from *D. sechellia* females is necessary to obtain a better understanding of the effects of *desatF* expression.

D. sechellia males court *D. simulans* females (Cobb and Jallon 1990) as found in these studies. The introgressions in *D. sechellia* females resulted in only one eloF NIL eliciting a higher CI from *D. sechellia* males than CI toward *D. simulans* females. This suggests that even if CHC structure is being altered in the NILs, the alteration is not sufficient to affect courtship. *D. sechellia* males are not inhibited by the lack of 7,11-HD (Cobb and Jallon 1990), and this is reflected in the similar CI of *D. sechellia* males toward both NIL and *D. simulans* females. The lack of *eloF* or *desatF* activity in the NILs, or of both genes in *D. simulans*, has only a minimal effect on the willingness of *D. sechellia* males to court.

Conclusions and Future Directions

Reproductive isolation is a mechanism that drives species divergence and maintains species barriers, helping to build the framework for evolutionary change (Dobzhansky 1935). Courtship between potential mating partners can be a step toward successful reproduction, thus changes to genes that affect courtship can reproductively isolate individuals from one another. Because behaviors involved in courtship are typically quantitative complex traits, it can be difficult to uncover specific genes contributing to those traits. My results suggest a small gene region around *eloF* is affecting reproductive isolation, presumably through an effect on CHCs,

although the functional role needs to be tested. These results support the idea that reproductive isolation can be caused by a relatively simple genetic basis (Coyne 1992).

I have shown the *eloF* introgression region to be a contributing factor in the pre-mating reproductive isolation between *D. simulans* and *D. sechellia*. Despite the complex nature of courtship behavior in *Drosophila*, I have provided evidence in support of single genes being sufficient to have a major effect on the willingness of a *D. simulans* male to court. My results expand on the role of *eloF* as it is already known in *D. melanogaster* (Chertemps et al. 2007), suggesting *eloF* expression changes may have played a large role in the evolution of some species in the *D. melanogaster* subgroup.

While *eloF* appears to be a major contributing factor to courtship between *D. simulans* and *D. sechellia*, QTL analysis of CHC production differences between the species implicates other regions of the genome as well (Gleason et al. 2009). Future work could expand on this study by examining genes found in different QTL regions, to further piece together the genetic interactions that work to affect CHC production and thus courtship as well. Advancements in technologies used to study individual genes, such as the CRISPR-Cas9 system (Yin et al. 2014), are making it increasingly easier and faster to manipulate the genome. The use of such technology to finely target and manipulate specific genes such as *desatF* and *eloF*, or other genes within QTL associated with this study, could help fully define the roles of these genes in influencing courtship. CHC analysis would also be useful in determining the functional roles of the genes, and would allow further elucidation of CHC biosynthesis pathways. Such work could lead to a much greater understanding of how the genetics of complex behaviors such as courtship work to influence reproductive isolation, and thus affect evolutionary change.

Literature Cited

Anholt, R. R. and T. F. Mackay. 2001. The genetic architecture of odor-guided behavior in *Drosophila melanogaster*. Behavioral Genetics 31:17-27.

Ashton, K., A. P. Wagoner, R. Carrillo, and G. Gibson. 2001. Quantitative trait loci for the monoamine-related traits heart rate and headless behavior in *Drosophila melanogaster*. Genetics 157:283-294.

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

Chertemps, T., L. Duportets, C. Labeur, R. Ueda, K. Takahashi, K. Saigo, and C. Wicker-Thomas. 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.

Chertemps, T., L. Duportets, C. Labeur, M. Ueyama, and C. Wicker-Thomas. 2006. A female-specific desaturase gene responsible for diene hydrocarbon biosynthesis and courtship behaviour in *Drosophila melanogaster*. Insect Molecular Biology 15:465-473.

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

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

Coyne, J. A. 1992. Genetics of sexual isolation in females of the *Drosophila simulans* species complex. Genetics Research 60:25-31.

Dobzhansky, T. 1935. A critique of the species concept in biology. Philosophy of Science 2:344-345.

Ferveur, J. F., F. Savarit, C. J. O'Kane, G. Sureau, R. J. Greenspan, and J. M. Jallon. 1997. Genetic feminization of pheromones and its behavioral consequences in *Drosophila* males. Science 276:1555-1558.

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

Gleason, J. M., K. A. Cropp, and R. S. Dewoody. 2004. DNA preparations from fly wings for molecular marker assisted crosses. *Drosophila* Information Service 87:107-108.

Gleason, J. M. and M. G. Ritchie. 2004. Do quantitative trait loci (QTL) for a courtship song difference between *Drosophila simulans* and *D. sechellia* coincide with candidate genes and intraspecific QTL? Genetics 166:1303-1311.

Gloor, G. E., W. 1992. Single-fly DNA preps for PCR. *Drosophila* Information Service 71:148-149.

Greenspan, R. J. and J. F. Ferveur. 2000. Courtship in *Drosophila*. Annual Review of Genetics 34:205-232.

Keurentjes, J. J., L. Bentsink, C. Alonso-Blanco, C. J. Hanhart, H. Blankestijn-De Vries, S. Effgen, D. Vreugdenhil, and M. Koornneef. 2007. Development of a near-isogenic line population of *Arabidopsis thaliana* and comparison of mapping power with a recombinant inbred line population. Genetics 175:891-905.

Krupp, J. J., C. Kent, J. C. Billeter, R. Azanchi, A. K. So, J. A. Schonfeld, B. P. Smith, C. Lucas, and J. D. Levine. 2008. Social experience modifies pheromone expression and mating behavior in male *Drosophila melanogaster*. Current Biology 18:1373-1383.

Legendre, A., X. X. Miao, J. L. Da Lage, and C. Wicker-Thomas. 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.

Lyman, R. F. and T. F. Mackay. 1998. Candidate quantitative trait loci and naturally occurring phenotypic variation for bristle number in *Drosophila melanogaster*: the *Delta-Hairless* gene region. Genetics 149:983-998.

Mackay, T. F. 2001. The genetic architecture of quantitative traits. Annual Review of Genetics 35:303-339.

Mas, F. and J. M. Jallon. 2005. Sexual isolation and cuticular hydrocarbon differences between *Drosophila santomea* and *Drosophila yakuba*. Journal of Chemical Ecology 31:2747-2752.

Robin, C., R. F. Lyman, A. D. Long, C. H. Langley, and T. F. Mackay. 2002. *hairy*: A quantitative trait locus for *Drosophila* sensory bristle number. Genetics 162:155-164.

Shirangi, T. R., H. D. Dufour, T. M. Williams, and S. B. Carroll. 2009. Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*. PLoS Biology 7 (8): e1000168.

Wang, L., X. Han, J. Mehren, M. Hiroi, J. C. Billeter, T. Miyamoto, H. Amrein, J. D. Levine, and D. J. Anderson. 2011. Hierarchical chemosensory regulation of male-male social interactions in *Drosophila*. Nature Neuroscience 14:757-762.

Wittkopp, P. J., B. K. Haerum, and A. G. Clark. 2004. Evolutionary changes in *cis* and *trans* gene regulation. Nature 430:85-88.

Yin, H., W. Xue, S. Chen, R. L. Bogorad, E. Benedetti, M. Grompe, V. Koteliansky, P. A. Sharp, T. Jacks, and D. G. Anderson. 2014. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nature Biotechnology 32:551-553.

Table 3.1. Genotyping primers and reaction conditions

Gene ID, name	Forward Primer (5' - 3')	Reverse Primer (5'- 3')	Annealing Temp.	PCR cycles	*D. simulans* size	*D. sechellia* size
Whole body/wing genotyping						
CG16905, *eloF*	ATTGCCATGCTGGCGATTTG	GACAGGATCCTCCGAAATGA	62°C	35	447	403
CG7923, *desatF*	AACTCATTCGATCGCCATTC	CGCATCAGATTCGTAAAGCA	62°C	35	437	323
***eloF* flanking markers**						
CG12806, teh1	CAGCCAACAGGAGTTAAAAGC	CACAAAGCCATCGAGTCAGT	60°C	35	139	153
CG9459 – CG16904 *	GCAGANNCGAAAGCCNTACA	ACAGTGTCCTTTCCGTGGAC	60°C	35	1650	869, 631, 150
CG8534+	GTCCGCCTGCCACTCATC	CTGCATAATGTTGTATGCCCTGA	60°C	35	701	386, 307
CG8516	GCGGATTCCCAAGTAAGTCC	AGTGCATCCAACTGGAACTG	60°C	35	156	150
CG9470, mtna	AACTCAATCAAGATGCCTTGC	TTGCAGTCAGATCCGCAGT	52°C	30	319	250
***desatF* flanking markers**						
CG7958, tna	ACGCCATGGAAAAGTCACTC	TGTGGGGCATCTGAAGTAGA	60°C	35	139	153
CG43693	GGAACATTTCGCAATGAATG	CGACTGACCGATAAATGCAA	60°C	35	494	516

PCR conditions: Initial denaturation at 95°C for 2 minutes followed by PCR cycles of 95°C for 1 minute, the annealing temperature for 1 minute, and extension at 72°C for 1 minute. After all cycles were completed, a further 72°C, 5 minute hold was done. *The forward primer is for CG9459 and the reverse primer is for CG16904. The PCR product was digested with the restriction enzyme DraI which cut the *D. sechellia* allele.

+CG8534 PCR product was digested with restriction enzyme TaqI, which cut the *D. sechellia* allele.

Table 3.2. RT-PCR primers and reaction conditions

Gene ID, name	Forward Primer (5' - 3')	Reverse Primer (5'- 3')	Annealing Temp.	PCR cycles
CG7923, *desatF*	GGTTCTTTGCCCACATAGGA	TTCGATGAACTTGGTGGTCA	57°C	30
CG16905, *eloF*	GTGGCGTCATCAGGGTCTAC	TTCAGCATGTACAAAATGCAAA	61°C	30
CG7939, *rp49*	CAGAATCTTATGACCATCCGCCCAGCA-TAC	CAGGAATTCAACGTTTACAAATGTGTAT-TC	65°C	30
CG15573, *Femcoat*	GGACCACAACTAATGCTGCTG	TCTTTCTTTTCGTCCCACCA	61°C	35
CG8982, A*cp26Aa*	GAACCTGATTTTGTTATGCTCTCA	TGGGAAGGAAGAGTGGAAGA	62°C	35

PCR conditions: Initial denaturation at 95°C for 2 minutes followed by PCR cycles of 95°C for 1 minute, the annealing temperature for 1 minute, and extension at 72°C for 1 minute. After all cycles were completed, a further 72°C, 5 minute hold was done.

Table 3.3. Courtship Latency and Courtship Index Overall

Male	Female	Courtship latency (seconds)[a]	Courtship index [a]
D. simulans	*D. simulans*	242.88 ± 116.03	0.558 ± 0.173[c]
	D. sechellia	470.30 ± 201.02	0.002 ± 0.002[d]
	NIL	314.20 ± 123.50	0.060 ± 0.045
	All females	342.46 ± 172.86[b]	0.207 ± 0.274
D. sechellia	*D. simulans*	742.28 ± 181.28	0.015 ± 0.006[d]
	D. sechellia	637.06 ± 280.16	0.129 ± 0.098[c]
	NIL	668.03 ± 270.57	0.086 ± 0.075
	All females	682.46 ± 237.37[b]	0.076 ± 0.082

[a] Courtship latency and courtship index are given as means and standard deviations for all pairings.

[b] CL toward 'All females' was lower from *D. simulans* males than from *D. sechellia* males (t-test $P<0.001$).

[c] CI of *D. simulans* males to *D. simulans* females was higher than CI of *D. sechellia* males to *D. sechellia* females (t-test $P<0.001$).

[d] CI of *D. sechellia* males to *D. simulans* females was higher than CI of *D. simulans* males to *D. sechellia* females (t-test $P=0.001$).

Figure 3.1

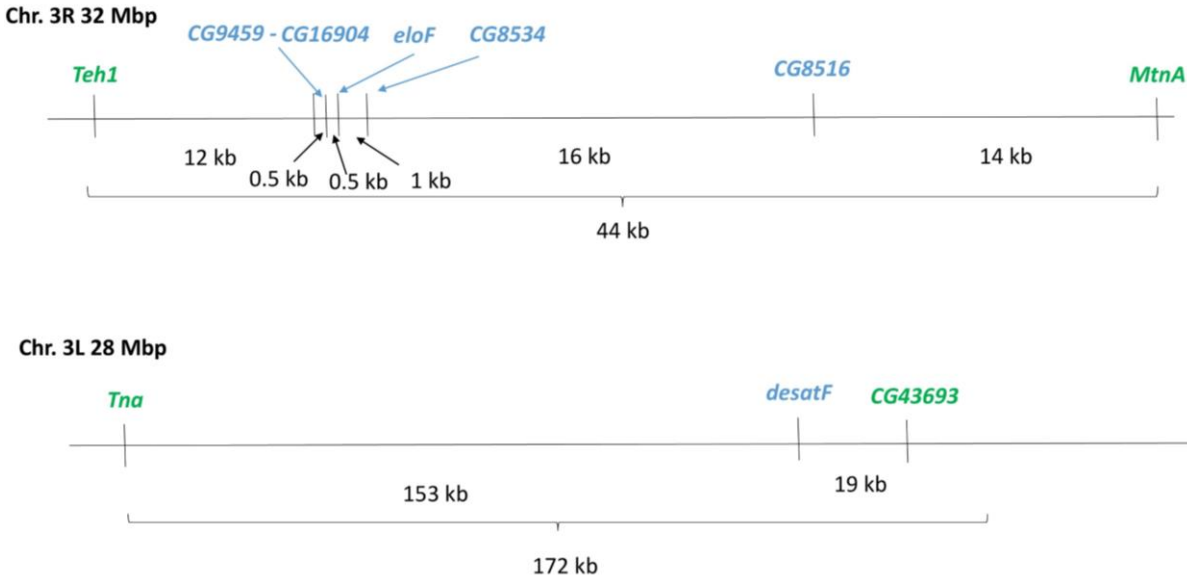
Chr. 3R 32 Mbp
CG9459 - CG16904
eloF
CG8534
Teh1
CG8516
MtnA
12 kb
0.5 kb
0.5 kb
1 kb
16 kb
14 kb
44 kb
Chr. 3L 28 Mbp
Tna
desatF
CG43693
153 kb
19 kb
172 kb

Figure 3.1. NIL introgression regions with genotyped markers. To determine the limits of each introgression, each NIL was genotyped for either the *D. simulans* (blue) or *D. sechellia* (green) allele of markers (Table 3.1) flanking *eloF* or *desatF*.

Figure 3.2

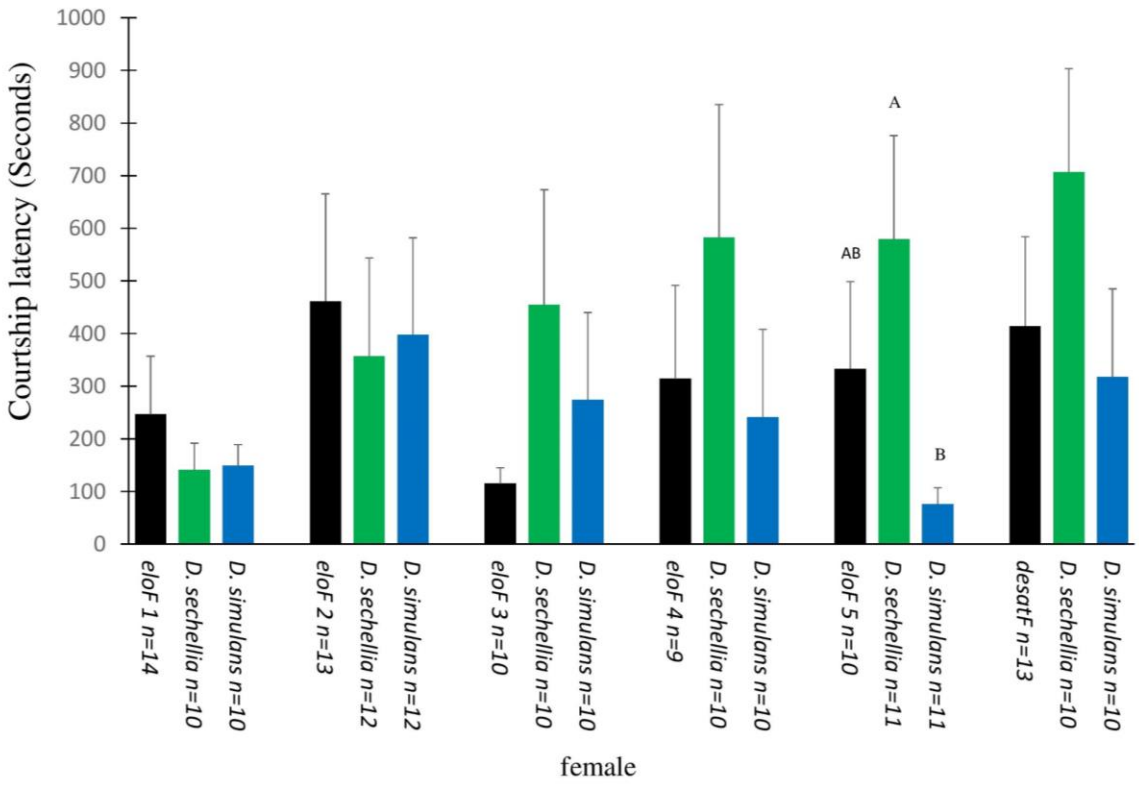
Courtship latency (Seconds)
1000
900
800
700
600
500
400
300
200
100
0
AB
A
B
eloF 1 n=14
D. sechellia n=10
D. simulans n=10
eloF 2 n=13
D. sechellia n=12
D. simulans n=12
eloF 3 n=10
D. sechellia n=10
D. simulans n=10
eloF 4 n=9
D. sechellia n=10
D. simulans n=10
eloF 5 n=10
D. sechellia n=11
D. simulans n=11
desatF n=13
D. sechellia n=10
D. simulans n=10
female

Figure 3.2 Courtship Latency from a *D. simulans* male does not differ among samples. In each trial, a *D. simulans* male was paired with a NIL female, a *D. simulans* female, and a *D. sechellia* female. Below the name of each female is the number of trials of each pairing. Error bars are standard error. Letters above bars indicate significant differences in each trial group (Student t-test, $P<0.05$).

Figure 3.3

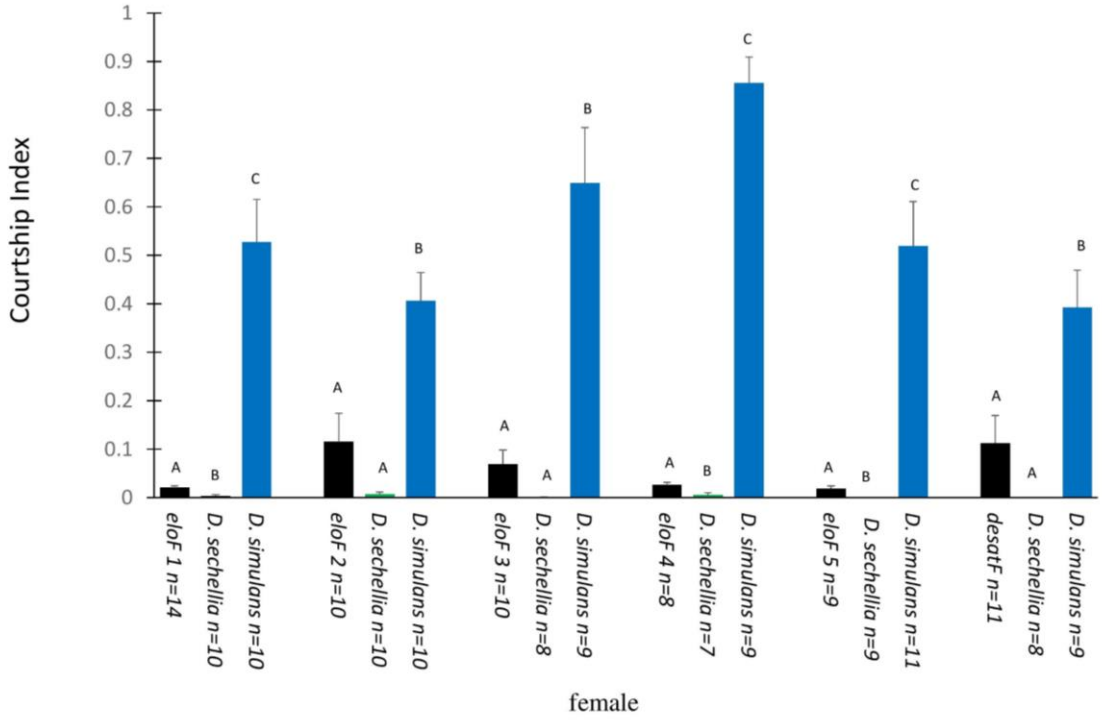
Courtship Index
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
A
B
C
A
A
B
A
A
B
A
B
C
A
B
C
A
A
B
eloF 1 n=14
D. sechellia n=10
D. simulans n=10
eloF 2 n=10
D. sechellia n=10
D. simulans n=10
eloF 3 n=10
D. sechellia n=8
D. simulans n=9
eloF 4 n=8
D. sechellia n=7
D. simulans n=9
eloF 5 n=9
D. sechellia n=9
D. simulans n=11
desatF n=11
D. sechellia n=8
D. simulans n=9
female

Figure 3.3 Courtship Index with a *D. simulans* male differs among NILs and *D. sechellia* female. In each trial, a *D. simulans* male was paired with a NIL female, a *D. simulans* female, and a *D. sechellia* female. Below the name of each female is the number of trials of each pairing. Error bars are standard error. Letters above bars indicate significant differences in each trial group (Student t-test, $P<0.05$).

Figure 3.4

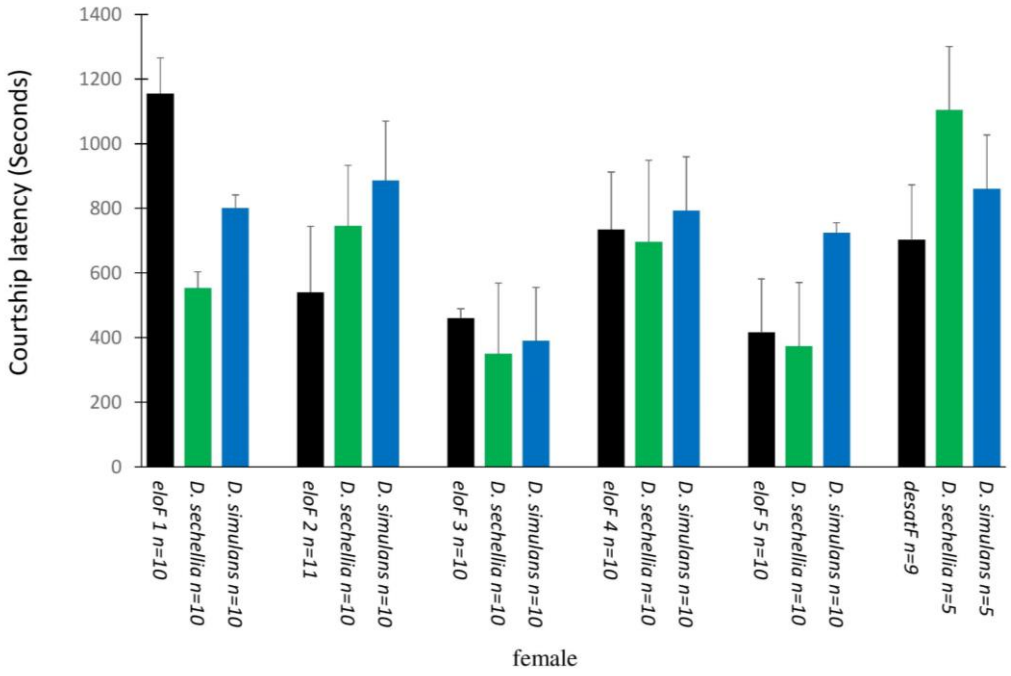
Courtship latency (Seconds)
1400
1200
1000
800
600
400
200
0
eloF 1 n=10
D. sechellia n=10
D. simulans n=10
eloF 2 n=11
D. sechellia n=10
D. simulans n=10
eloF 3 n=10
D. sechellia n=10
D. simulans n=10
eloF 4 n=10
D. sechellia n=10
D. simulans n=10
eloF 5 n=10
D. sechellia n=10
D. simulans n=10
desatF n=9
D. sechellia n=5
D. simulans n=5
female

Figure 3.4. Courtship Latency from a *D. sechellia* male does not differ among samples. In each trial, a *D. sechellia* male was paired with a NIL female, a *D. simulans* female, and a *D. sechellia* female. Below the name of each female is the number of trials of each pairing. Error bars are standard error. Letters above bars indicate significant differences in each trial group (Student t-test, $P<0.05$).

Figure 3.5

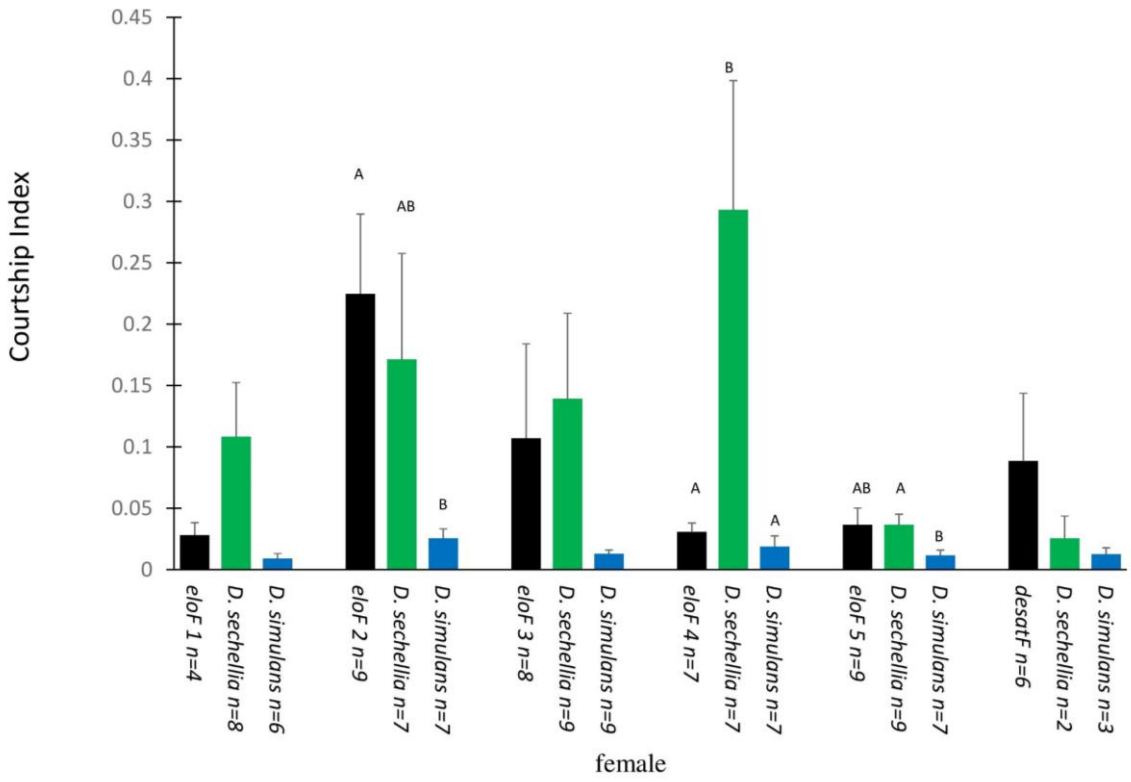
Courtship Index
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
A
AB
B
A
B
A
AB
A
B
eloF 1 n=4
D. sechellia n=8
D. simulans n=6
eloF 2 n=9
D. sechellia n=7
D. simulans n=7
eloF 3 n=8
D. sechellia n=9
D. simulans n=9
eloF 4 n=7
D. sechellia n=7
D. simulans n=7
eloF 5 n=9
D. sechellia n=9
D. simulans n=7
desatF n=6
D. sechellia n=2
D. simulans n=3
female

Figure 3.5. Courtship Index with *D. sechellia* male differs among some lines. In each trial, a *D. sechellia* male was paired with a NIL female, a *D. simulans* female, and a *D. sechellia* female. Below the name of each female is the number of trials of each pairing. Error bars are standard error. Letters above bars indicate significant differences in each trial group (Student t-test, $P<0.05$).