2011

THE GENETIC BASIS OF INTERSPECIFIC FEMALE PREFERENCE IN DROSOPHILA SIMULANS

Meghan Elizabeth Laturney

THE GENETIC BASIS OF INTERSPECIFIC FEMALE PREFERENCE IN *DROSOPHILA SIMULANS*

(Spine title: Genetics of Interspecific Female Preference in *Drosophila*)

(Thesis format: Integrated-Article)

by

Meghan E. Laturney

Graduate Program in Biology

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

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Supervisor	Examiners
Dr. Amanda Moehring	Dr. Elizabeth MacDougall-Shackleton
Supervisory Committee	Dr. Anthony Percival-Smith
Dr. Graham Thompson	Dr. Klaus-Peter Ossenkopp
Dr. Anthony Percival-Smith	

The thesis by

Meghan Elizabeth Laturney

entitled:

The Genetic Basis of Interspecific Female Preference in *Drosophila simulans*

is accepted in partial fulfillment of the
requirements for the degree of
Master of Science

Date: Thursday December 22, 2011

Chair of the Thesis Examination Board

Abstract

Behavioral differences between species have a genetic basis and contribute to species isolation. Genomic regions have been identified that influence female rejection of heterospecific males on the right arm of the third chromosome in *Drosophila*, however, no individual loci have been identified. Here, I used deficiency mapping to locate regions that influence the rejection of *D. melanogaster* males by *D. simulans* females. First, I tested two genes that have been previously shown to affect female within-species mate choice: neither of these genes was found to contribute to between-species female preference. Next, I identified five small significant regions that contain candidate genes contributing to behavioral isolation, which were all located in areas of low interspecific recombination. Furthermore, I identified the first candidate gene for behavioural isolation in *Drosophila*. I also provided a list of candidate gene. Identification of genes that influence behavioural isolation will provide understanding of the genetic influence on biodiversity.

Keywords: speciation, reproductive isolation, behavioral isolation, deficiency mapping, *Drosophila melanogaster*, *Drosophila simulans*.

Acknowledgements

First and foremost I would like to express my gratitude to my supervisor Dr. Amanda Moehring for the hours of dedicated work she gave to this project and to the development of me as a scientist. I would like to thank her for always making time for my questions about the project, or research in general, and providing a comfortable environment to learn. Her work ethic, creativity, intelligence, problem solving skills, and above all her desire for discovery will always serve as an incredible example for me. I would also like to thank Dr. Graham Thompson for his help with the development of this project and his influence in the writing process. Furthermore, as an instructor, he has forced me to think differently about the process of evolution. I would particularly like to thank Jessica Pardy for her assistance with the maintenance of the lines and crosses, and her hours spent collecting in the fly room. Her time and effort was never without appreciation, and I could not have completed this without her help. I would like to thank the entire Moehring lab for their contribution to this project, especially Helene LeVasseur-Viens for her continuous help, Rachelle Kanippayoor for her help with the dissections, and Craig MacNair for scoring some of the behavior data. Finally I would like to thank Brent Sinclair for providing the wild-type *D. melanogaster* lines, and to Kevin Cook and Annette Parks at the Bloomington Drosophila Stock Center for their quick and comprehensive responses related to the deficiency lines questions.

Table of Contents

List of Tables

List of Figures

List of Appendices

Glossary

Alternative splicing: The processes in which the components of the gene sequence that are transcribed are assembled in different ways to produce different gene products.

Asymmetrical behavioural isolation: The phenomenon in which there is a different rate of heterospecific mating between males of species A with females B compared to males of species B with females of species A.

Background genetic effect: The gene or genes in the genome that are not under investigation that may influence a trait

Basal isolation the behavioural isolation between two allopatric species before populations are reintroduced causing an increase in behavioural isolation.

Co-dapted gene complex: Genes inherited together (due to tight linkage or genes found within an inversion) that are selected for or against as a unit instead of each gene individually.

Conspecific an individual belonging to the same species

Conspecific sperm precedence the phenomena that is observed when a female mates with both conspecific and heterospecific males, allowing for their gametes to compete, and the majority of offspring are produced from the conspecific mating.

Courtship latency the time elapsed from the introduction of a male and female to the commencement of the male courtship.

Crossability the ease in which heterospecific mating occurs.

Deficiency breakpoints the locations on the chromosome that the defines the region that is deleted (two breakpoints per deletion).

Deficiency lines stocks of *Drosophila melanogaster* that have two alleles for each locus except for a specific region that has only one allele per loci due to a deletion on one of the homologous chromosomes.

Explosive speciation the divergence of one species into a large number of new species in a relatively short amount of time and usually resulting in a large degree of variation appearance, behaviour, and ecological niche specialization.

Female choosiness the phenomena that is observed when females mate significantly more with conspecific than heterospecific males and can vary depending on the species pair.

Gal4-UAS system a system that can change gene expression to test the effect of a specific gene on a trait of interest. The gene product of Gal4 interacts with a sequence of DNA called the upstream activation sequence (UAS) and increases the transcription of the gene sequence that directly follows it. To assess the increased expression of the gene on a trait, the trait is observed in UAS/gene of interest flies and Gal4+UAS/gene of interest flies.

Genetic basis an umbrella term referring to the parts of the genome that influences a trait. This includes regions of the genome, specific genes, alleles, variants and all types of interactions between these elements.

Genetic hitchhiking an allele can increase in frequency not due to selection for it but due to selection for an allele that is in close proximity in the genome. Due to the

physical linkage, recombination does not break the alleles apart. Therefore, as the selected allele increases in frequency so too will the other linked alleles.

Hybrid lethality hybrids are a result of heterospecific matings and contain one set of homologous chromosomes for each species. In some cases, the interaction of two different genomes within the same individual causes one or both sexes of hybrids to be inviable.

Hybrid sterility hybrids are a result of heterospecific matings and contain one set of homologous chromosomes for each species. In some cases, the interaction of two different genomes within the same individual causes one or both sexes of hybrids to be sterile.

Hybrid speciation a speciation event which results from a heterospecific mating and the production of a hybrid. Hybrids then assortatively mate creating a new species.

Hybridization the event in which a heterospecific mating occurs and hybrids are produced.

Homogametic mating a form of assortative mating, This includes mating between two individuals with relatively similar genomes. This similarity can refer to individuals from the same species, with the same inversion polymorphism, or with the same set of alleles.

Introgression study a study that uses introgression lines to map a trait. When mapping species-specific traits, introgression lines are created by crossing a male and a female from different species (for example species A and B, respectively). F_1 hybrids females are then crossed with males of one of the parental species (for

example, species A). The F_2 generation is then crossed again with the same parental species (species A) and this is repeated for many generations. The result is individuals with an almost entirely species A genome with small regions of species B and these regions are identified with use of molecular markers. These individuals are then tested for species-specific traits and if the trait resembles species B, then the regions of the genome that originated from species B may be causing the species-specific trait.

Interaction effect an effect that is observed with considering at least two independent variables (IV) and one dependent variable (DV) and the two IVs both influence the trait in a non-additive way.

Interline comparison a comparison between different deficiency lines to reduce the area that may contain gene(s) for the trait of interest. For example, if line A was significant, then a gene or genes for the trait of interest may be within the deleted region. However, if line B overlapped line A and was not significant, then the region considered to contain gene(s) for the trait of interest would be reduced to only include the area unique to line A.

Interspecific chromosomal inversion polymorphism one element of structural differences of the genome of closely related species and results in the order of the genes along the chromosome to appear in a different order. A chromosomal inversion is a rearrangement of genetic material in which a segment of a chromosome is broken and inserted in a reversed fashion so that if the original chromosome was ordered ABCD the new arrangement could read ACBD. Therefore,

interspecific chromosomal inversion polymorphisms cause a species-typical gene order.

Loss of function mutation a change in the genetic material that results in the gene product having reduced or no function.

Multimodal signalling a complex signal that is composed of multiple elements.

Paracentric inversion a chromosomal rearrangement that results from a region that does not contain the centromere breaking off and inserted back in the opposite direction.

Pleiotropic the phenomenon of a single gene influencing more than one phenotypic traits.

Recombination mapping the type of mapping that uses introgression lines to map a trait.

Reinforcement the increased behavioural isolation between two species in sympatric populations.

Sexual maturation speed the rate at which an organism reaches sexual maturity.

Sexual selection the process in which attractive mates produce more offspring than other individuals in the population and therefore alleles for such traits increase in the population. These traits may or may not be an honest signal or genetic quality.

Speciation the process in which one interbreeding population diverges into two or more isolated interbreeding populations with new independent evolutionary lineages.

Visible marker a mutant phenotype that indicates the presence of a dominant allele in the genome. Visible markers can indicate which homologous chromosome was inherited.

List of Abbreviations

bal: balancer chromosome

CHC: Cuticular hydrocarbons

Df: deficiency chromosome

IPI: interpulse interval

mel: *Drosophila melanogaster*

oe-: Flies without working oenocytes

sim: *Drosophila simulans*

QTL: Quantitative trait loci

X: X chromosome

XR: right arm of the X chromosome

2L: left arm of the second chromosome

2R: right arm of the second chromosome

3L: left arm of the third chromosome

3R: right arm of the third chromosome

7,11-HD: *cis, cis* 7,11-heptacosadience

7-T: 7-trisosene

Chapter 1: General Introduction

With the massive amount of biodiversity on the planet, an obvious question is how are new species formed and maintained? This query has fuelled research aimed at understanding the process of speciation and species isolation. From this body of work, three camps representing the main mechanisms of species isolation have emerged: pre-zygotic isolation, post-copulatory prezygotic isolation, and post-zygotic isolation. The latter of the three has received the most attention which may be due to the relative ease of scoring the two barriers (e.g. hybrid inviability and sterility) of this mechanism. This work has uncovered some of the genetic machinery behind the causes of interspecific inviability (Brideau *et al.* 2006) and interspecific sterility (Mihola *et al.* 2009). Post-copulatory prezygotic mechanisms, such as conspecific sperm precedence, have also received some attention and have been found to be wide-spread, isolating species of both animals (Price 1997) and plants (Arnold *et al.* 1993). Lastly, prezygotic isolating barriers are those that stop gene flow between two interbreeding populations before the formation of a zygote and are present regardless of whether species are geographically separated or not. In other words, both sympatric species and species that have not come into contact can exhibit behavioural isolation. Of the three mechanisms prezygotic barriers are thought to be the first to evolve during the process of speciation (Coyne and Orr 1997). Therefore, understanding behavioural isolation not only can help to answer how species boundaries are maintained but also how are they formed.

One prezygotic barrier is behavioural isolation, which usually results from closely related species having divergent mating signals: one or both of the sexes fails to identify the other as a suitable mate (Blows and Allan 1998; Coyne 1992; Doi *et al.* 2001; Moulin *et al.* 2004). Therefore, the types of behaviours that typically separate species are the male trait and/or the female preference for or against that trait. Although mating behaviours have been shown to isolate species as females mate more readily with conspecific than heterospecific males (Moerhing *et al.* 2004; Moehring *et al.* 2006) they can be very difficult to study because of numerous sources of environmental and genetic variation that influence the trait (Gefen and Gibbs 2009; Narraway *et al.* 2010). In order to understand the variation in mating behaviour that is seen between species, each source of variation is examined independently. For example, by rearing and observing organisms in a uniform "common garden" environment, we can study the effect of genetic variation on this behaviour. By identifying the genetic variants that cause interspecific differences in mating behaviour, we can determine which mutations and alterations in the genetic material cause the differences in behaviour between two isolated populations. Yet, little is known about the individual genetic variants that contribute to behavioural isolation.

Mating behaviour is a quantitative trait that is most likely influenced by multiple genes (Moehring and Mackay 2004). Despite its importance for speciation, knowledge of the genetic basis of behavioural isolation is not well understood. The reason for this is primarily attributable to a two-fold dilemma. First, the established methods in genetics for locating genes that contribute to a trait, namely

recombination mapping, necessitate crossing two divergent lines and producing fertile offspring. However, by definition, separate species usually do not produce either fertile or viable offspring. Second, identifying the genetic basis of a behaviour (a quantitative trait) is complicated. It requires the location of multiple genes with different effect sizes, necessitating a repeatable measure of the behaviour, large sample sizes, and the availability of powerful genetic tools such as a sequenced genome (Anholt and Mackay 2004).

Despite these obstacles, the genetic basis of behavioural isolation has been studied in different species of animals and plants. For example, the genetic basis of floral scent production in *Petunia axillaris* (Petunia) has been examined, which plays an important role in pollinator attraction and thus contributes to isolation between related species of plants (Klahre *et al.* 2011). However, most of the advances in this field have been within the insect and fish model systems.

Butterflies are an attractive model system to use for investigating behavioural isolation due to the relationship consistently seen between wing colour and mate preference, the relative ease of measuring the external colour phenotype (Morehouse and Rutowski 2010). For example, in *Heliconius* butterflies the mating of two differently patterned species can create hybrid offspring with a unique pattern unlike either parent. Furthermore, the offspring displays a preference for the hybrid pattern over both of the parental colour patterns suggesting a possibility for hybrid speciation (Melo *et al.* 2009). This association between the forewing colour trait and preference exists because it is caused by the same gene (*wingless*) or multiple genes close linked to *wingless* (Kronforst *et al.* 2006).

Cichlid fish are another system that has advanced our knowledge of the genetic basis of behavioural isolation. The cichlids in Lake Victoria have been used as a traditional example of explosive speciation and have highlighted the importance of behavioural isolation in speciation as there are no post-zygotic mechanisms for isolation among these cichlid species (Maan *et al.* 2004). Strong directional intraspecific selection driven by female choice on male colouration has caused a great divergence of species-specific male colouration, and thus has played a role in behavioural isolation (Maan *et al.* 2004). Furthermore, interspecific female mating preference for conspecific colouration is heritable with only a few loci responsible (Haesler and Seehausen 2005). Preference for male colouration and display, however, is not the entire story. Non-visual cues such as odour have also been implicated in interspecific behavioural isolation among cichlids (Blias *et al.* 2009) and host-parasite coevolution may have contributed to the divergence in the major histocompatibility complex (MHC), causing changes in the sensory systems such as olfaction, which in turn influences mate choice between cichlid species (Blias *et al.* 2009).

Although butterfly colouration/preference coupling, and cichlid behavioural isolation due to colouration and odour have provided insight into the genetic basis of behavioural isolation, these systems are limited in that they do not have the powerful genetic tools available in the model system of *Drosophila*. The impact of *Drosophila* on this area of research has been so profound since both of the primary obstacles to speciation research, outlined above (two-fold dilemma), can be more easily overcome compared to the butterfly and cichlid systems. First, many

Drosophila sister species are only partially isolated in a lab setting, producing viable and fertile hybrids (Coyne 1992). Second, *Drosophila melanogaster* has a large number of genetic tools available (Matthews *et al.* 2005), including a sequenced genome, readily-available gene mutant lines, and systems to control gene expression, which allows for easy identification and manipulation of genes.

The genetic basis of *Drosophila* mating behaviour

The mating behaviour of *D. melanogaster* follows a series of steps, involves many stereotypic endophenotypes, and contains multimodal signaling that can be executed stereotypically even if a fly has developed in isolation (Hall 1994; Greenspan 1995; Greenspan and Ferveur 2000). Courtship behaviour progresses from the orientation of the male to the female, the male tapping the female's abdomen with his front leg (exchange of chemical signals), the male vibrating to produce a species-specific courtship song (auditory signal), licking the female's genitalia (exchange of chemical signals), and finally attempting copulation by curling his abdomen towards the female (Hall 1994; Greenspan 1995; Greenspan and Ferveur 2000). If she is receptive, the female will slow her locomotion, orient herself toward the male, present her abdomen, and spread her genital plates to allow him to copulate (Griffith and Ejima 2009). If she is unreceptive, she will extrude her ovipositor making copulation physicially impossible, increase her locomotion, or fly away. If the male is unsuccessful at initiating copulation, he will return to an earlier stage of courtship and begin the process again. The duration of copulation, once achieved, is species-specific, with *D. melanogaster*'s copulation averaging approximately 20 minutes (Fowler 1973).

The genetic basis of *Drosophila* mating behaviour was first investigated by mutagenesis studies, whereby individual genes were mutated and their effects on behaviour were assessed. These studies resulted in the identification of many genes that are important to the formation of normal mating behaviour of *D. melanogaster*. Some examples of mutations that influence male mating behaviour are *period* (Kyriacou and Hall 1980), *fruitless* (Gailey and Hall 1989), *cacophony* (Kulkarni *et al.* 1988), *slowpoke* (Peixoto and Hall 1998), and *nerd* (Ferveur and Jallon 1993). The only mutants found to influence female receptivity are *dissatisfaction* (Finley *et al.* 1997), *spinster* (Suzuki *et al.* 1997), and *chaste* (Juni and Yamamoto 2009). These studies were of great importance as they provided crucial information into both the sensory system used in *Drosophila* mating and the types of genes that can influence the construction of mating behaviour.

Mutagenesis studies helped to shed light on what sensory signals males and females use to assess potential mates during courtship. Males must select suitable individuals to court as to not waste efforts on inappropriate endeavours (Dickson 2008). However, mutations in *voila*, a gene expressed in gustatory receptors in the foreleg of the male fly, caused an increase in male-male courtship (Balakireva *et al.* 1998). Without the normal functioning of these receptors, males could not inhibit improper courtship. This suggests that males use gustation to perceive male-typical compounds which cause them to avoid courtship attempts with other males. Females, on the other hand, must determine whether to accept or reject a courting male. Flies have non-volatile compound on their cuticle which act as pheromones (Ferveur 2005). Mutations in *nerd*, which produced males with significantly less

male-typical compounds, were consistently rejected by females even with vigorous courtship attempts. Similarly, the uni-lateral wing vibrations that produce a species-specific song is a hallmark of male *Drosophila* courtship. Mutations in such genes as *period* (Kyriacou and Hall 1980) and *cacophony* (Kulkarni *et al.* 1988) produced males that generated abnormal song and had reduced mating success. This suggests that both the correct male pheromone profile (Ferveur and Jallon 1996) and courtship song are important for female receptivity.

Mutagenesis studies have also revealed the types of genes that can influence the construction of normal *Drosophila* mating behaviour. Mutations that altered ion channel functions and disrupted normal locomotor activity also influenced the ability of males to produce a normal courtship song (*slowpoke*, *paralytic*, and *no action potential*; Peixoto and Hall 1998). Mutations in genes that code for a transcription factor influenced both male behaviour (*fruitless*, Hall 1978), and female receptivity (*dissatisfaction*, Finley *et al.* 1998). Finally, these studies showed the genes for female receptivity (*spinster*, *dissatisfaction*, and *chaste*) along with multiple genes for male mating behaviour (*voila*, *fruitless*, *5-HT7*) are expressed in the central nervous system (Suzuki *et al.* 1997; Finley *et al.* 1997, 1998; Juni and Yamamoto 2009; Balakireva *et al.* 1998; Hall 1978; Becnel *et al.* 2011). Although mutagenesis studies have been greatly useful for understanding the genetic basis of within species mating behaviour, they may not be as useful in determining the genetic basis of between species mating behaviour due to their focus on within-species mating, failure to test allelic variation, and exclusive focus on male traits.

The first limitation that accompanies the mutation studies when attempting to apply them to questions of species divergence is that they have focused on identifying the genetic basis of intraspecific (within species) mating behaviour. Although important, these results do not necessarily shed light on how new species are formed since the genes that control intraspecific behavioural interactions may or may not be the same genes that are important in interspecfic behavioural interactions. Is mate recognition a continuous range varying from assortative mating within a species to the rejection of heterospecific males? Blows and Allan (1998) argued that if species isolation was produced by sexual selection, then the traits involved in species isolation should be the same traits used in both sexes during within-species mate choice.

To test this hypothesis, Blows and Allan (1998) performed a series of crosses between *D. serrata* and *D. birchii*, which have overlapping geographic ranges along the east coast of Australia. Although morphologically very similar, there is strong behavioural isolation and weak post-zygotic isolation (Ayala 1965) between the two species. The crosses created hybrid lines, and these lines were used to determine if hybrids use species-specific characters to select appropriate mates. Blows and Allan showed there are typical cuticular hydrocarbon (CHC) profiles, which are used as pheromones, for different sexes and different species. With the use of perfuming experiments, these researchers determined that the same systems (CHC and olfaction) are used in within-species mate choice (sexual selection) and between-species mate choice (behavioural isolation). This shows that the same variation in the same trait can be used for both within- and between-species mate choice. For

example, genes for olfactory system development could therefore be important for normal female mating behaviour in both species. However, the genetic basis for the interpretation of the CHC profile may vary between species and may not be captured in the mutagenesis studies. Although the two species use the same sensory system for mate choice, it is not clear if the genetic basis of this trait or preference for this trait is the same in the different species: the loci that determine within-species attractiveness may not be the same loci that determine between-species unattractiveness.

Investigation into this question led to a series of studies that investigated the relationship between interspecific hybridization and intraspecific receptivity. Carracedo *et al.* (1989) proposed that genes controlling interspecific mating behaviour could be pleiotropic for other phenotypes, such as female intraspecific receptivity. They proposed that if females are slower to accept conspecific males they may also be more reluctant to accept a heterospecific males, which would be observed as lower hybridization with *D. simulans* males and thus higher behavioural isolation. This could eventually link the genetic basis of both behaviours. In a natural population, a high level of within-species receptivity would be selected against due to its pleitropic effect of high interspecies hybridization. Females that mated readily with conspecific males but also with heterospecific males would sire less viable offspring and would contribute less genetic information into the next generation than females that exclusively mated with conspecifics. Experiments in the lab as well as natural experiments such as island population helped to shed light on the validity of the relationship between intra- and interspecific mating behaviour.

In the lab, heterospecific mating was selected for in females by collecting the conspecific offspring from females that mated readily with heterospecific males. After multiple generations, females not only had shorter copulation latency (time to start copulation) with heterospecific males but also conspecific males. Therefore, interspecific and intraspecific female receptivity may be controlled by the same genes (Pineiro *et al.* 1993). To directly test the impact of increased interspecific mating on intraspecific mating, another study selected for hybridization and after 12 generations of selection hybridization levels increased from 10% to 79%. When these females were placed in a choice assay with a conspecific and heterospecific fly of the opposite sex, almost no heterospecific matings were observed. In nature, multiple mates are available, and therefore, selection for heterospecific mating is unlikely to influence mate choice (Izquierdo *et al.* 1992).

In island populations, when migrants populate a new island it is likely that the least choosy females will propagate the most offspring (Kaneshiro 1983). The most choosy females may not find a high quality male, and therefore, will not generate any offspring. Assuming that low intraspecific choosiness results in high hybridization rates, we expect isolated island species to have high levels of hybridization. Surprisingly, the opposite trend is observed: females from island populations such as *D. mauritiana* are more choosy against males from their ancestral species *D. simulans* than females of the ancestral species against males of the island population. In other words, *D. mauritiana* (island) females do not mate readily with *D. simulans* (ancestral) males, but *D. simulans* females mate readily with *D. mauritiana* males (Moehring *et al.* 2004).

Finally, recombination mapping studies have located regions of the genome that influence behavioural isolation, which do not include the genes identified through mutagenesis (Moehring *et al.* 2004, 2006). Therefore, genes that specifically influence behavioural isolation may not influence within species mating.

The second limitation of mutagenesis studies is that they eliminate the gene's function in order to test whether it affects a behaviour. While this shows us that the gene is important for creation of the behaviour, it does not necessarily tell us anything about the naturally-occurring genetic variation that contributes to the differences seen between species. Yet, it is these variants that evolution can act upon to cause shifts in behaviour within a species, or isolation between species.

Many studies have investigated if genes identified through mutagenesis studies contribute to variation in mating behaviour. The genes important for normal male mating behaviour were not found to contribute to variation in courtship behaviour (Carney 2007), did not contribute to variation between low and high mating male lines (Moehring and Mackay 2004), and did not vary in expression in a natural population of *D. melanogaster* (Ruedi and Hughes 2009). The genes important for normal female mating behaviour were also not found to vary in expression between courted and naïve same age virgin females (Lawniczak and Begun 2004). The genes identified through mutagenesis consistently do not appear to influence the variation in mating behaviour within species, and therefore, may also not contribute to the variation between species (Mackay et al. 2005).

The final limitation of mutagenesis studies is that most studies on mating behaviour in *Drosophila* have focused on the male aspects of mating. Limiting the

view to just male mating behaviour ignorantly implicates the male as the only active individual in the courtship/copulation process. This bias is likely due to the relative ease of scoring the male courtship behaviour compared to the female rejection behaviour. Although males preferentially court conspecific females with larger body sizes, which is a good indicator of female fecundity (Byrne and Rice 2006), and in some species (for example *D. virilis*), males are able to discriminate against heterospecific females (Nickel and Civetta 2009), it is ultimately up to the female if copulation occurs in most *Drosophila* species (Greenspan 1995). Females will more often discriminate against heterospecifics than males do (Moehring *et al.* 2004), and can easily prevent unwanted copulations by flying away from the courting male or extruding her ovipositor (Hall 1994). Therefore, when studying the genetic basis of behavioural isolation, more attention should be given to genes for female preference.

Mutagenesis studies provided proof that genes can influence mating behaviour, insight into the systems that *Drosophila* use during courtship, and revealed the types of genes that influence the construction of behaviour. However, further investigation into the genetic basis of mating behaviour showed that most of the genes identified do not influence species-specific behaviour, did not contribute to variation in mating behaviour, and have grossly overlooked the importance of female receptivity. Genes identified through mutagenesis may be crucial to the normal development of the trait but may not vary between species.

The modes of *Drosophila* male signalling during courtship

To date, we have no strong candidate genes implicated in interspecific female receptivity (Butlin and Ritchie 2001; Kyriacou 2002). The variability we see in female preference, both within and between species, is most likely dictated by the integration of the auditory and olfactory systems (Yamamoto and Nakano 1999). To complicate investigation of these two systems, the amount that females of each species rely on one system over the other is most likely species-specific (Doi *et al.* 2001; Moulin *et al.* 2004; Tomaru *et al.* 2004; Gleason *et al.* 2005). A gene for behavioural isolation, specifically one for interspecific female preference, is most likely going to be associated with the signalling pathway of the auditory system (Wheeler *et al.* 1991; Doi *et al.* 2001) used to recognize differences in male courtship song characteristics, the olfactory system used to recognize CHC pheromone profiles (Blows and Allan 1998), or both systems in the organization of the part of the brain that receives and interprets signals from both pathways (Yamamoto and Nakano 1999; Dickson 2008). This is because both modes of signalling are used during *Drosophila* courtship (Hall 1994; Greenspan 1995; Greenspan and Ferveur 2000) and both male signals vary between species (Blows and Allan 1998; Doi *et al.* 2001; Civetta and Cantor 2003; Gleason *et al.* 2005). A candidate region for such integration is the mushroom body which receives signals from many sensory systems in *Drosophila* (Davis 2011) including the olfaction system (de Belle and Heisenberg 1994), and has been linked to sexual behaviour (O'dell *et al.* 1995, Balakireva *et al.* 1998) specifically female receptivity (Neckameyer 1998). If gene(s) for female receptivity is largely dependent on the types of signalling used during

courtship, the similarities in the genetic basis of interspecific mating may extend only as far as the similarities in the sensory system reach (Etges 2002). Therefore, a gene for interspecific female preference in *D. melanogaster* may be specific to the species, to the *D. melanogaster* group, or to *Drosophila* in general but may not influence mate choice in other insects, invertebrates, or other types of organisms.

Courtship song

There are two main elements to the *Drosophila* courtship song (the sine song and the interpulse interval), and males of different species usually differ from each other on both aspects (Kyriacou 2002). A female's ability to identify conspecific song over heterospecifics can lead to behavioural isolation (Doi *et al.* 2001). For females in the melanogaster group, the most important element of courtship song is the interpulse interval (IPI) which differs among the males, and preference for variants of IPI seems to differ among females (Gleason 2005). The most famous gene to influence courtship song is the *period* (*per*) gene. Mutations in this gene influence IPI (Kyriacou and Hall 1980) and transgenic *D. melanogaster* flies with *D. simulans per* produced *D. simulans* typical rhythm (Wheeler *et al.* 1991). Instead of a species difference reflecting a complex genetic basis, the species differences in song rhythm reflect just a small number of amino acid changes (Wheeler *et al.* 1991). Females from this same transgenic line showed associated preference for the transgenic male's IPI (Ritchie and Kyriacou 1994), and a later study also showed evidence of assortative mating with a different *per*-transgenic line (Sakai and Ishida 2001).

Although the genetic basis of this preference is not straightforward, it is clear that females may be using variation of song between species in mate choice. Females

can detect male song and male movement with use of the receptors in the antenna and neurons from here project to the dorsal brain, which requires feminization in order for females to be receptive (for review, see Dickson 2008). Likewise, many genes that regulate the sex determination pathway have also been shown to affect courtship song, and thus potentially species-specific song which would convey the male's species identity (Gleason 2005).

Pheromones

Each species of *Drosophila* has characteristic cuticular hydrocarbons (CHCs) that act as a protective barrier to desiccation and most likely evolved as an adaptation to dry climates (Rouault *et al.* 2004). CHCs are nonvolatile compounds that are detected by both males and females, most likely through touch (gustation) at close proximity rather than smell at long distances (Ferveur 2005). Detection of the CHC profile occurs through a large family of odorant receptors that send information about the environment via odorant sensory neurons to the antennal lobe, which is analogous to the olfactory bulb in mammals (for a review, see Dickson 2008).

In addition to dessication resistance, CHCs have also been shown to be important in mating behaviour (Ferveur 2005) and used during mate selection as pheromones (Billiter *et al.* 2009, Coyne *et al.* 1994). Billeter *et al.* (2009) used a Gal4-UAS system to block the development of oenocyetes (cells specialized to produce the cuticular hydrocarbons). Flies without working oenocytes (oe-) were completely devoid of all CHCs but behave normally. However, behaviour towards oe- individuals was significantly altered: oe- females received more courtship than wild-

type females, and oe- males were courted and received attempted copulations by wild-type males. Therefore, this study showed that CHCs may act as a sexual identity signal.

Altering the CHC pheromone profile through containment methods can break down behavioural isolation barriers between species by increasing courtship towards heterospecific females that have been altered so that their CHC profile resembles that of conspecific females (Coyne *et al.* 1994). Therefore, hydrocarbon profiles contribute to the sexual isolation between these species. Normally, *D. simulans* males do not court wild-type *D. melanogaster* females, however, they did court oe- *D. melanogaster* females. Re-application of a *D. melanogaster* female CHC onto the oe- females suppressed this courtship (Billeter *et al.* 2009). Therefore, CHCs can act as both sexual and species identification.

Although there are more than 20 different CHC molecules on the cuticle of the fly, only the predominant hydrocarbons have received much examination (Gleason *et al.* 2009). *D. simulans* and *D. mauritiana* have a monomorphic CHC profile, with the main hydrocarbon of both males and females being the same 23-carbon chain compound, *cis* 7-trisosene (7-T). However, *D. melanogaster* and *D. sechellia* are dimorphic, with the males having large amounts of 7-T but females lacking this hydrocarbon and instead have large amounts of a 27-carbon molecule, *cis, cis* 7,11-heptacosadience (7,11-HD; Ferveur *et al.* 1996).

Through mutagenesis studies, genes have been identified to affect CHC production, such as *dsat1* and *dsat2* (Dallerac *et al.* 2000), *Enhancer of zest* (Wicker-Thomas and Jallon 2001), *Ddc* (Marican *et al.* 2004), *nerd* (Ferveur and Jallon 1993),

seven pentacosene and *smoq* (Ferveur and Jallon 1996), as well as some sex determination genes, such as *doublesex* (Jallon *et al.* 1988). However, only the genetic basis of the main CHC components (7-T and 7,11-HD) have been examined. Additionally, it is unclear if variation in these genes produces the variation that is seen in CHC production between populations of the same species, or variation in production between species (Coyne *et al.* 1999; Ritchie and Noor 2004; Gleason *et al.* 2005).

Interspecific differences have been investigated in an attempt to understand the genetic basis of behavioural isolation. *D. simulans* males do not readily court female *D. sechellia* but do court *D. simulans*/*D. sechellia* hybrids (Coyne *et al.* 1994). *D. simulans* has a monomorphic CHC profile and the difference that exists between these two species is thought to affect their behavioural isolation (Gleason *et al.* 2005). The chromosomal basis of variation in 7-T between these two species, and thus the divergence of the main component of their pheromone profiles, is primarily caused by a single locus, or multiple closely linked loci, on the third chromosome explaining both variation in mating behaviour and pheromone production (Civetta and Cantor 2003). In 2005, Gleason *et al.* identified four quantitative trait loci (QTLs) that control the differential production of 7-T and 7,11-HD in *D. simulans* and *D. sechellia* females. For 7-T production, one QTL was found on the right arm of the X chromosome (XR in *D. melanogaster*), and three on the right arm of the third (3R in *D. melanogaster*). For 7,11-HD production, two QTLs were found on the third chromosome. Epistatic interactions among loci were also found, which implies that each locus may produce machinery for one step in a multi-step process of

hydrocarbon production. In 2009, Gleason *et al.* found more evidence to support the earlier QTLs, and also found four additional CHCs that differ between these females when they explored the entire CHC profile.

What drives the change in female preference for different CHC profiles? When populations become isolated, the separated groups likely occupy two different environments, which may have different selective pressures causing different allelic frequencies. Over time, new and different mutations can occur in the two isolated populations, which the different selective pressures can act upon. Therefore, it is possible that different CHC profiles are more advantageous in different environments and this may have been the cause of the variation in CHC profiles observed between species. However, there is more to it than just natural selection. In order for CHC profiles to change over time, female preferences must change in parallel at the population level.

Evidence for this hypothesis comes from a study that isolated populations of *D. serrata* and reared them for multiple generations on different media (Rundle *et al.* 2005). They found that both the CHC profile of the flies and the preference for it in the females differed between the populations. The change in environment along with the genetic variability within the genome caused a change in the CHC profiles which was mirrored in the female preference for the change. The specific genetic changes that caused the differences between the two groups, however, remain unknown.

From the research dedicated to identifying the genetic basis of CHC variation between species and courtship song variation between males of different species we

can comfortably deduce that different species have different CHC profiles, different courtship songs, and females preferentially mate with conspecific males based at least partially on each signal. Through mutagenesis studies and QTL mapping, genes responsible for CHC and song production and differences between species have been partially realized. However, as mate choice goes the signal is only half of the story: an investigation into the preference for the trait must be done.

To date, no individual genes have been identified as influencing intra- or inter-specific female preference in *Drosophila*, although the trait has a clear heritable basis (Hall 1994). Unfortunately, the majority of studies seeking to address this question have been done in *Drosophila* species other than *D. melanogaster* (Noor *et al.* 2001; Doi *et al.* 2001; Moehring *et al.* 2006; Coyne 1992; Civetta and Cantor 2003; Moehring *et al.* 2004), which do not have the genetic tools available for further refinement. Although the genome of 12 different species of *Drosophila* have been sequenced (*Drosophila* 12 Genomes Consortium 2007) it is only recently that tools such as Gal4-UAS system to manipulate gene expression and transposon vectors for use in mutagenesis studies available for *D. melanogaster* have been modified for the different species of *Drosophila* (Holtzman *et al.* 2010). A few studies have made use of the behavioural isolation between *D. melanogaster* and *D. simulans*, but have done so in a very limited way so that whole chromosome arms containing thousands of genes have been identified to influence behavioural isolation, but not specific gene variants (Carracedo *et al.* 1995; Uenoyama and Inoue 1995; Carracedo *et al.* 1998b). Despite the shortcomings, various genomic regions have been identified that contribute to behavioural isolation in multiple species of *Drosophila*.

Genetic basis of species isolation for different species pairs

D. pseudoobscura and *D. persimilis*

D. pseudoobscura are found across much of Western North America, and are located both in sympatry and in allopatry with *D. persimilis*. The initial genetic basis of isolation between these species, termed basal isolation, was found to be caused by only two regions in the genome: one on the left arm of the X chromosome (which is homologous to the X in *D. melanogaster*) and one on the second chromosome (homologous to the right arm of chromosome 3, called 3R, in *D. melanogaster*), within an interspecific inversion that differentiates *D. pseudoobscura* and *D. persimilis* (Noor *et al.* 2001).

Female *D. pseudoobscura* from sympatric regions hybridize less with male *D. persimilis* than females from allopatric regions without *D. persimilis* (Noor 1995). The phenomenon of females from sympatric populations displaying greater behavioural isolation than females from allopatric populations is referred to as reinforcement. Ortiz-Barrientos *et al.* (2004) investigated the genetic basis of the increased discrimination of sympatric *D. pseudoobscura* females. By introgressing (crossing) pieces of the sympatric *D. pseudoobscura* genome, into an allopatric *D. pseudoobscura* background they mapped the increase in behavioural isolation to two alleles of strong effect, one on the right arm of the X chromosome (called *Coy-1*; translates to 3L in *D. melanogaster*) and one on the fourth chromosome (called *Coy-2*; translates to 2L in *D. melanogaster*). However, Barnwell and Noor (2008) used six pairs of different inbred strains in a QTL study to try to replicate the previous findings. They could not, and therefore determined that *Coy-1* and *Coy-2*, although

they may be important, are not the only loci causing increased behavioural isolation in sympatric *vs.* allopatric populations. These alleles may be present at low frequencies in natural populations and present in all inbred lines.

Each of the *D. pseudoobscura* sympatric and allopatric *Coy2* alleles were introgressed into a *D. persimilis* background (creating per*Coy2*sym and per*Coy2*allo lines; Ortiz-Barrientos and Noor 2005). If the reinforced behavioural isolation was caused by an increased receptivity for *D. pseudoobscura* (conspecifics) by the *D. pseudoobscura* sympatric population, the expected results would be that per*Coy2*sym females are more likely to mate with *D. pseudoobscura* than per*Coy2*allo, but instead they found the opposite: per*Coy2*sym females were less likely to mate with *D. pseudoobscura* than per*Coy*2allo. This suggests that an allele for reduced interspecific mating within a species (*Coy2*sym) can cause the same reduction in interspecific mating when placed within another species (Ortiz-Barrientos and Noor 2005).

The explanation provided by Ortiz-Barrientos and Noor is that *Coy-2* may be a One-Allele mating locus. This theory suggests that one allele (*Coy-2*) can exist in both the sympatric population of *D. pseudoobscura* and in *D. persimilis* population, and aids in the reinforced behavioural isolation between these populations, but not in the basal behavioural isolation. In other words, the same allele causes females of both species to have an increased discrimination against heterospecifics. This is possible if, for example, the gene encodes for increased odour sensitivity or reduced dispersal (Servedio and Noor 2003). This theory would explain why per*Coy2*sym females were less likely to mate with *D. pseudoobscura* than per*Coy2*allo.

An alternative explanation is that the QTL that supports the reinforced behavioural isolation is the same locus (or closely linked to a gene) that causes assortative mating between the allopatric and sympatric populations of *D. pseudoobscura*. Isolation between the sympatric and allopatric populations of *D. pseudoobscura* has never been tested, but if there is assortative mating between these two populations, it is possible that *Coy-2* could be responsible as it has been shown to influence behavioural differerences between the two populations. This could only be the case if the focal *D. pseudoobscura* male used in the behaviour assays was from the allopatric population, and thus, the sympatric *D. pseudoobscura* females mated less compared to the allopatric females (showing disassortative mating).

D. ananassae and *D. pallidosa*

D. ananassae and *D. pallidosa* are present in overlapping geographic regions. Males of both species court females of both species, but there is strong interspecific female preference that reduces the gene flow between the two. The genetic basis of this behaviour was first explored with female F_1 hybrids and the hybrids were found to prefer *D. ananassae* males over *D. pallidosa* males (Doi *et al.* 2001). This suggests that *D. ananassae* genes for interspecific female choice must be dominant over those from *D. pallidosa*. The same study created introgression lines to locate the genomic regions responsible for this behaviour. A region on the left arm of the second chromosome (homologous to 3R in *D. melanogaster*) near the *Delta* locus was identified to play a role in female species mate choice: females that were almost entirely *D. pallidosa* except for a small region near the delta locus mated significantly

more with *D. ananassae* males and significantly less with *D. pallidosa* males (Doi *et al.* 2001). In other words, this locus both increased intraspecific mating in *D. ananassae* and decreased interspecific mating between *D. ananassae* females with heterospecific males. This region was later confirmed by a study that found 2L (3R in *D. melanogaster*) to be important for the willingness of *D. pallidosa* females to mate with *D. ananassae* males; and XL, 2L, and 3R (X, 3R, and 2L in *D. melanogaster*, respectively) for *D. ananassae* female's willingness to mate with *D. pallidosa* males. All of the identified regions had species specific inversions (Sawamura *et al.* 2008), suggesting that regions of the genome with reduced recombination between the species may be more likely to harbour behavioural isolation loci.

D. santomea and *D. yakuba*

D. santomea and *D. yakuba* diverged approximately 400 thousand years ago (Llopart *et al.* 2002). *D. yakuba* population is wide-spread across Africa including some of the islands off the coast. On one of these islands *D. santomea* are found and although this species pair has a small overlapping geographic region, no reinforcement has been observed (Lachaise *et al.* 2000). Male courtship behaviour may contribute to the behavioural isolation between these two species as *D. santomea* males do not court heterospecific females with any vigour. To investigate the genetic basis behind the female interspecific mating, Moehring *et al.* (2006) created a QTL map for female rejection of heterospecific males. For *D. santomea* female discrimination against *D. yakuba* males, Moehring and colleagues detected three QTLs: two on the X chromosome (homologous to X in *D. melanogaster*), and one on the third chromosome (3R in *D. melanogaster*).

D. simulans and *D. sechellia*

D. simulans is a cosmopolitan species, while its closely-related sibling species *D. sechellia* is only found on the Seychelles Islands in the Indian Ocean. Coyne (1992) found asymmetrical behavioural isolation between *D. simulans* and *D. sechellia* in that *D. simulans* females are less choosy against *D. sechellia* males than are *D. sechellia* females against *D. simulans* males. Hybrids have an intermediate level when paired with *D. simulans* males, suggesting an additive genetic basis for interspecific female preference. Coyne (1992) used backcrossing to locate the chromosomes that are important for this behaviour in both species. This technique was not possible for *D. sechelia* because females with small regions of *D. sechelia* in an otherwise *D. simulans* background showed readily mated with *D. simulans* males. However, the genetic basis in *D. simulans* was localized to both the second and third chromosomes (2 and 3 in *D. melanogaster*) which had a moderate and strong effect, respectively.

D. simulans and *D. mauritiana*

D. simulans is a cosmopolitan species and *D. mauritiana* is only found on the island of Mauritius in the Indian Ocean. It is thought that *D. mauritiana* resulted from colonization by a recent common ancestor with *D. simulans* about 250 000 years ago (Kliman *et al.* 2000). Females of these species are almost identical, and the males are only distinguishable by the shape of their genital arch (Tsacas and David 1974). Asymmetrical species isolation is present, with *D. simulans* being the less choosy of the two courted females. The absence of mating is due to the refusal of females since females of both species are courted vigorously by males of both species (Coyne 1989). Hybrids produced by *D. mauritiana* males and *D. simulans* females mate

readily with *D. simulans* males, and thus, the genes for interspecific mate discrimination in *D. mauritiana* must be recessive (Coyne 1989, 1992). By backcrossing the hybrids to *D. mauritiana* males, Coyne was able to measure each *D. mauritiana* chromosome's effect on decreasing mating with *D. simulans* males. He found each of the main autosomes have very large effects with the effect of X being very small (Coyne 1989). Further dissections of the second chromosome determined that each arm of the second chromosome contains at least one gene for reducing *D. mauritiana* female matings with *D. simulans* males (2R and 2L in *D. melanogaster*; this method of uncovering recessive *mauritiana* genes also possibly removed *D. simulans* genes for conspecific mate preference – these genes may or may not be one in the same). Moehring *et al.* (2004) examined the same pairings but with a more refined map and found seven QTLs that contribute to *D. mauritiana* discrimination against *D. simulans* males: two on the X chromosome (X in *D. melanogaster*), two on the second chromosome (2 in *D. melanogaster*), and three on the third chromosome (3 in *D. melanogaster*). Although *D. simulans* females are not choosy and readily mate with *D. mauritiana* males, matings between these two species are abnormally short and result in no or limited sperm transfer, decreasing the number of hybrid offspring (Coyne 1992).

D. simulans and *D. melanogaster*

D. melanogaster and *D. simulans* have overlapping geographic ranges. Although both females show behavioural isolation, *D. simulans* females are far more choosy against *D. melanogaster* males (Wood *et al.* 1980; Carracedo *et al.* 2003): interspecific crosses with *D. melanogaster* females are produced with relative ease

in the lab but the reciprocal cross very rarely occurs (Carracedo *et al.* 1998a). F_1 hybrids made from *D. melanogaster* females are all sterile females, and from the reciprocal cross are all sterile males. Due to this sterility, the conventional method of QTL mapping is not possible as this would require backcrossing to one of the parental species. Therefore, other methods used to determine the genetic basis of behavioural isolation between these two species were employed.

A genomic region was identified on the third chromosome for *D. melanogaster* female receptivity, and genomic regions on all three major chromosomes were identified for rejection of *D. simulans* males by *D. melanogaster* females (Carracedo *et al.* 1995). Although there is some evidence that male *D. simulans* may contribute to the behavioural isolation (Carracedo *et al.* 2003), there is no such evidence for discrimination by *D. melanogaster* males (Jamart *et al.* 1993). Therefore, the strong behavioural isolation demonstrated by *D. simulans* females is largely due to rejection of heterospecific (*D. melanogaster*) males.

To investigate this, different lab strains of *D. simulans* females (Carracedo *et al.* 1998b; Carracedo *et al.* 2000) and *D. melanogaster* males (Carracedo *et al.* 1998b) were compared for the rate of interspecific mating. Crossability varied between both stains of *D. melanogaster* males and strains of *D. simulans* females (Uenoyama and Inoue 1995; Carracedo *et al.* 1998b) but were still highly correlated (Carracedo *et al.* 2000). When strains of *D. simulans* were crossed, the pure species F_1 females were then crossed *D. melanogaster* males and the crossability was compared to the two parental strains. Mixed results were found: while one study found that F_1 females always showed greater levels of hybridization (Uenoyama and Inoue 1995), another

study found that in most cases F_1 females showed significantly lower levels of hybridization (Carracedo *et al.* 1998b). Further inconsistencies include one study that found that X and the third chromosome act additively to contribute to the rejection of *D. melanogaster* males by *D. simulans* females (Uenoyama and Inoue 1995), while another study found that the X and the left arm of the second chromosome influenced the trait (Carracedo *et al.* 1998a). These results may be due to the low genetic variability within inbred laboratory lines, or they may support the hypothesis that the genetic basis of behavioural isolation may be different for each species, and even different between populations of the same species.

M and Z forms of *D. melanogaster*

D. melanogaster are found all over world, usually commensally with humans, and it was once thought that there was gene flow between populations, including those found spread across large continents (Kreitman and Aguade 1986). However, a Zimbabwe population was found to have twice the amount of genetic variation compared to North American populations, with certain variants only present in Zimbabwe (Begun and Aquadro 1993). Females from these Zimbabwe lines (Z) show strong behavioural isolation against males from cosmopolitan regions (M): when they have the choice, Z females prefer to mate with Z males, but show no post-mating isolation (hybrid sterility or inviability) when they are mated with M males. Females from cosmopolitain regions also show behavioural isolation with Z males, but it is weaker than that seen in Z females (Wu *et al.* 1995). The genetic basis for this strong preference in Z females was mapped to all three major chromosomes, with the largest effect being contributed by the third chromosome (Hollocher *et al.*

1997). With the use of recombinant lines and visible markers (dominant mutations to identify which homologous chromosome was inherited from which parental species), Ting *et al.* (2001) identified the genetic basis of the female preference in Z females for Z males. They identified a region of large effect, and another locus or loci of minor effect on the left arm of the third chromosome (3L), as well as a region on the right arm (3R), which most likely houses two loci.

Conclusions

Various genomic regions varying from whole chromosomes, chromosomal arms, sub-chromosomal regions, to specific QTLs have been identified to contribute to the behavioural isolation between species pairs (reviewed above). Although the genetic basis may be species-pair-specific (Carracedo *et al.* 2000), one common attribute of these loci is their location in the genome: most of these loci fall within species inversion-polymorphisms or near the centromere or telomere. Regions that influence behavioural isolation between *D. santomea* and *D. yakuba* were found near the centromere on 3R (Moehring *et al.* 2006), and near the telomere for the *D. simulans* and *D. mauritiana* species pair (Moehring *et al.* 2004) and M and Z form assortative mating in *D. melanogaster* (Ting *et al.* 2001). Loci responsible for the behavioural isolatation between *D. ananassae* and *D. pallidosa* (Sawamura *et al.* 2008), and the isolation between *D. pseudoobscura* and *D. persimilis* (Noor *et al.* 2001) all fell within interspecific inversion polymorphisms. However, this was not true for the regions responsible for increased behavioural isolation caused by reinforcement in the later species pair (Ortiz-Barrientos *et al.* 2005). However, these

loci for reinforcement have not been confirmed by follow up studies (Barnwell and Noor 2008).

Inversions have also been shown to play a role in within-species assortative mating. Unlike other species of *Drosophila*, *D. ananassae* males have spontaneous meiotic recombination which contributes to the entire species having a high degree of inversion polymorphisms. One inversion, alpha, is a large paracentric inversion covering the majority of 2L (3R in D. *melanogaster*). To investigate whether this inversion could contribute to behavioural isolation within this species, Nanda and Singh (2001) created karyotypically different strains homozygous for one of three naturally occurring inversions. Through mate choice assays, they found a preference for homogamic matings in all three populations.

Although there is sequence divergence between the species of *Drosophila* that have been sequenced, most genes are orthologs, and the structure of the genome is well conserved with approximately similar number of protein coding genes and minimal gene shuffling, however, inversions do exist between different species (Drosophila 12 Genome Consortium 2007). Similarly, areas of low recombination like those near the centromere and telomere have been implicated in reproductive isolation (Weetman *et al.* 2011).

Genomic rearrangements, centromereic, or telomeric areas can act as an island of low recombination between the two populations and create and maintain gene complexes (genes inherited together). Over time, new mutations can occur within theses complexes, and due to reduced recombination (Stevison *et al.* 2011), and can create a population-typical phenotype if the complexes contain variants for

local adaptation (Feder *et al.* 2011). Therefore, even in the face of gene flow between the two groups a new population identity can be created. If mutations occur within these regions that cause a change in female preference by influencing assortative mating within species (Nanda and Singh 2001), these areas can encourage a speciation event and influence behavioural isolation between species (Lowry and Willis 2010; Noor *et al.* 2001).

The genetics of female preference is the missing piece to the puzzle for understanding the genetic basis of speciation, yet no individual gene has been identified to influence either within-species or between-species female preference in *Drosophila*. The main setback to these studies is the lack of tools available for identifying speciation genes. *D. melanogaster*, which has by far the greatest number of genetic tools of all of the *Drosophila* species, cannot be used in QTL studies due to the inability to produce F_1 fertile hybrids with any of its sibling species. Therefore, these studies are done in other *Drosophila* species that lack refinement tools like deficiency lines and readily available gene mutants. Without these, the genetic regions identified in previous studies cannot be further explored and specific genes for interspecific female preference will remain anonymous. Presently, the identification of a gene for behavioural isolation will only be identified if another method of gene mapping is identified to utilize the tools in *D. melanogaster* or another model organism.

In this thesis, I outline and perform a mapping technique that utilizes the tools available in *D. melanogaster* to locate the genetic basis for behavioural isolation between *D. melanogaster* and *D. simulans*. Within this species pair *D. simulans*

females are far more choosy against *D. melanogaster* males than are *D. melanogaster* females against *D. simulans* males. Since n *D. melanogaster* males readily court heterospecific females, the strong isolation between *D. simulans* females and *D. melanogaster* males is most likely due to the female's rejection behaviour. Therefore, this thesis attempts to identify the genetic basis of the interspecific female preference within *D. simulans* against *D. melanogaster* males.

Female hybrids created by crossing male *D. simulans* and female *D. melanogaster*, although sterile, are viable. These hybrids mate with *D. melanogaster* males suggesting that the *D. simulans* genes for heterospecific rejection are recessive to the *D. melanogaster* genes for conspecific receptivity (Davis *et al.* 1996). One mapping technique that maps for recessive genes with a large effect size and does not require fertile hybrids is deficiency mapping. I compared the mating propensity of female hybrids that inherited a small deficiency uncovering the *D. simulans* genome to hybrids that have a full diploid genome by placing them in a mating assay with *D. melanogaster* males, I will locate regions that may contains genes for behavioural isolation. The null hypothesis is that there are no genes for interspecific female preference in each region that is independently tested. This hypothesis will only be rejected if females that inherited the deficiency mate with *D. melanogaster* males significantly more than the diploid hybrid females. In this case, the alternative hypothesis that there are gene(s) within this region for interspecific female preference will be accepted. By systematically testing deficiencies spanning different regions of the right arm of the third chromosome, I have located genomic regions that contribute to behavioural isolation.

In chapter two, deficiencies that span the right arm of the third chromosome were tested, providing a broad map of the regions that may influence behavioural isolation. In chapter three, the regions that were identified in chapter two were finer mapped and the genes in these refined regions were investigated. The identification of relatively small genomic regions that contain genes for behavioural isolation within this species pair was the first step in my discovery of the first candidate gene responsible for interspecific female preference.

References

Anholt, R. R, and T. F. Macka. 2004. Quantitative genetic analysis of complex behaviours in *Drosophila*. 2004. Nat. Rev, Genet. 5:838-849.

Arnold, M. L., J. L. Hamrick, and B. D. Bennett. 1993. Interspecific pollen competition and reproductive isolation in Iris. J. Hered. 84:13-16.

Ayala, R. J. 1965. Relative fitness of populations of *Drosophila serrata* and *Drosophila birchii*. Genetics 51:527-544.

Balakireva, M., R. F. Stocker, N. Gendre, and J. F. Ferveur. 1998. *Voila*, a new *Drosophila* courtship variant that affects the nervious system: behavioral, neural, and genetic characterizations. J. Neurosci. 18:4335-4343.

Barnwell, C. V., and M. A. Noor. 2008. Failure to replicate two mate preference QTLs across multiple strains of *Drosophila pseudoobscura*. J. Hered. 99:653-656.

Becnel, J., O. Johnson, J. Luo, D. R. Nassel, C. D. Nichols. 2011. The serotonin 5-HT7Dro receptor is expressed in the brain of *Drosophila*, and is essential for normal courtship and mating. PLoS ONE 6:e20800.

Begun, D. J., and C. F. 1993. Aquadro. African and North American populations of *Drosophila melanogaster* are very different at the DNA level. Nature 365:548-550.

Billiter, 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.

Blias, J., M. Plenderleith, C. Rico, M. I. Taylor, O. Seehausen, C. van Oosterhout, and G. F. Turner. 2009. Assortative mating amoung Lake Malawi cichlid fish

populations is not simply predictable from male nuptial colour. BMC Evol. Biol. 9:53.

Blows, M. W., and R. A. Allan. 1998. Levels of mate recognition within and between two *Drosophila* species and their hybrids. Am. Nat. 152:826-837.

Brideau, N. J., H. A. Flores, J. Wang, S. Maheshwari, X Wang, and D. A. Barbash. 2006. Two Dobzhansky_muller genes interact to cause hybrid lethality in *Drosophila*. Science 314:1292-1295.

Byrne, P. G., and W. R. Rice. 2006. Evidence for adaptive male mate choice in fruit fly *Drosophila melanogaster*. Proc. Biol. Sci. 273:917-922.

Butlin, R., and M. G. Ritchie. 2001. Searching for speciation genes. Nature 412:31-32.

Carney, G. E. 2007. A rapid genome-wide response to *Drosophila melanogaster* social interactions. BMC Genomics 8.

Carracedo, M. C., A. Asenjo, and P. Casares. 1998a. Inheritance mode of *Drosophila simulans* female mating propensity with *D. melanogaster* males. J. Hered. 89:101-104.

Carracedo, M. C., A. Asenjo, and P. Casares. 2003. Genetics of *Drosophila simulans* male mating discrimination in crosses with *D. melanogaster*. Heredity 91:202-207.

Carracedo, M. C., L. Garcia-Florez, and E. San Miguel. 1989. Sexual maturation in *Drosophila melanogaster* females and hybridization with *D. simulans* males: a study of inheritance modes. J. Hered. 80:157-158.

Carracedo, M. C., R. Pineiro, and P. Casares. 1995. Chromosomal substitution analysis of receptivity and sexual isolating in *Drosophila melanogaster*. Heredity 75:541-546.

Carracedo, M. C., A. Suarez, A. Asenjo, and P. Casares. 1998b. Genetics of hybridization between *Drosophila simulans* females and *D. melanogaster* males. Heredity 80:17-24.

Carracedo, M. C., A. Suarez, and P. Casares. 2000. Sexual isolation between *Drosophila melanogaster*, *D. simulans* and *D. mauritiana*: sex and species specific discrimination. Genetica 108:155-162.

Civetta, A., and E. J. Cantor. 2003. The genetics of mating recognition between *Drosophila simulans* and *D. sechellia*. Genet. Res. 82:117-126.

Coyne, J. A. 1989. Genetics of sexual isolation between two sibling species, *Drosophila simulans* and *Drosophila mauritiana*. Proc. Natl. Acad. Sci USA 86:5464-5468.

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

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., and H. A. Orr. 1997. Patterns of speciation in *Drosophila* revisited. Evolution 51:295-303.

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

Dallerac, R., C Labeur, J. M. Jallon, D. C. Knipple, W. L. Roelofs, and C. Wicker-Thomas. 2000. A delta 9 desaturase gene with a different substrate specifically is responsible for the cuticular diene hydrocarbon polymorphism in *Drosophila* melanogaster. Proc. Natl. Acad. Sci. USA 15: 9449-9454.

Davis, R. L. 2011. Traces of *Drosophila* memeory. Neuron 14:8-19.

Davis, A. W., J. Roote, T. Morley, K. Sawamura, S. Herrmann, and M. Ashburner. 1996. Rescue of hybrid sterility in crosses between *D. melanogaster* and *D. simulans*. Nature 380:157-159.

de Belle, J. S., and M. Heisenberg. 1994. Associative odor learning in *Drosophila* abolished by chemical ablation of mushroom bodies. Science 263:692-695.

Dickson, B. J. 2008. Wired for sex: the neurobiology of *Drosophila* mating decisions. Science 322:904-909.

Doi, M., M. Matsuda, M. Tomaru, H. Matsubayashi, and Y. Oguma. 2001. A locus for female discrimination behavior causing sexual isolation in *Drosophila*. Proc. Natl. Acad. Sci. USA 98:6714-6719.

Drosophila 12 Genomes Consortium. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. Nature 450:203-218.

Etges, W. J. 2002. Divergence in mate choice systems: does evolution play by the rules? Genetica 116:151-166

Feder, J. L., R. Geiji, T. H. Powell, and P. Nosil. 2011. Adaptive chromosomal divergence driven by mixed geographic mode of evolution. Evolution 65:2157-2170.

Ferveur, J. F. 2005. Cuticular hydrocarbons: their evolution and roles in *Drosophila* pheromonal communication. Behav. Genet. 35:279-295.

Ferveur, J. F., M. Cobb, H. Boukella, and J. M. Jallon. 1996. World-wide variation in *Drosophila melanogaster* sex pheromone: behavioural effects, genetic bases and potential evolutionary consequences. Genetica 97: 73-80.

Ferveur, J. F., and J. M. Jallon. 1993. *Nerd*, a locus on chromosome II, affects male reproductive behavior in *Drosophila melanogaster*. Naturwissenschaften 80:474-475.

Ferveur, J. F., and J. M. Jallon. 1996. Genetic control of male cuticular hydrocarbons in *Drosophila melanogaster*. Genet. Res. 67:211-218.

Finley, K. D., P. T. Edden, M. Foss, E. Gross, N. Ghbeish, R. H. Palmer, B. J. Taylor, and M. McKeown. 1998. *dissatisfaction* encodes a tailless-like nuclear receptor expressed in a subset of CNS neurons controlling *Drosophila* sexual behavior. Neuron 21:1363-1374.

Finley, K. D., B. J. Taylor, M. Milstein, and M. McKeown. 1997. *dissatisfaction*, a gene involved in sex-specific behavior and neural development of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 94:913-918.

Fowler, G. L. 1973. Some aspects of the reproductive biology of *Drosophila*: sperm transfer, sperm storage, and sperm utilization. Adv. Genetic. 17:293-360.

Gailey D. A., and J. C. Hall. 1989. Behavior and cytogenetics of *fruitless* in *Drosophila melanogaster*: different courtship defects caused by separate closely linked lesions. Genetics 121:773-785.

Gefen, E., and A. G. Gibbs. 2009. Interactions between environmental stress and male mating success may enhance evolutionary divergence of stress-resistant *Drosophila* populations. Evolution 63:1653-1659.

Gleason, J. M. 2005. Mutations and natural genetic variation in the courtship song of *Drosophila*. Behav. Genet. 35:265-277.

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. 103:416-424.

Greenspan, R. J. 1995. Understanding the genetic construction of behavior. Sci. Am. 271:74-79.

Greenspan, R. J., and J. M. Ferveur. 2000. Courtship in *Drosophila*. Annu. Rev. Genet. 34:205-232.

Griffith, L. C., and A. Ejima. 2009. Multimodal sensory integration of courtship stimulating cues in *Drosophila melanogaster*. Ann. N. Y. Acad. Sci. 1179:394-398.

Haesler, M. P., and O. Seehausen. 2005. Inheritance of female mating preference in a sympatric sibling species pair of Lake Victoria cichlids: implications for speciation. Proc. Biol. Sci. 272:237-245.

Hall, J. C. 1978. Courtship among males due to a male-sterile mutation in *Drosophila melanogaster*. Behav. Genet. 8:125-141.

Hall, J. C. 1994. The mating fly. Science 264:1702-1714.

Hollocher, H., C. T. Ting, M. L. Wu, and C. I. Wu. 1997. Incipient speciation by sexual isolation in *Drosophila melanogaster*: extensive genetic divergence without reinforcement. Genetics 147:1191-1201.

Holtzman, S., D. Miller, R. Eisman, H. Kuwayama, T. Niimi, and T. Kaufman. 2010. Transgenic tools for members of the genus *Drosophila* with sequences genomes. Fly 4:349-362.

Izquierdo, J. I., M. C. Carracedo, R. Pineiro, and P. Casares. 1992. Response to selection for increased hybridization between *Drosophila melanogaster* females and *D. simulans* males. J. Hered. 83:100-104.

Jallon, J. M., G. Lauge, L. Orssaud, and C. Antony. 1988. Female pheromones in *Drosophila melanogaster* are controlled by the *doublesex* locus. Genet. Res. 51:17-22.

Jamart, J. A., M. C. Carracedo, and P. Casares. 1993. Sexual isolation between *Drosophila melanogaster* females and *D. simulans* males. Male mating propensities versus success in hybridization. Experientia 49:596-598.

Juni, N., and D. Yamamoto. 2009. Genetic analysis of *chaste*, a new mutation of *Drosophila melanogaster* characterized by extremely low female sexual receptivity. J. Neurogenet. 23:329-340.

Kaneshiro, K. Y. 1983. Sexual selection and the direction of evolution in the biosystematics of the Hawaiian Drosophilidae. Ann. Rev. Entomol. 28:161-178.

Kilman, R. M., P. Andofatto, J. A. Coyne, F. Depaulis, M. Kreitman, A. J. Berry, J. McCarter, J. Wakeley, and J. Hey. 2000. The population genetics of the origin and divergence of the *Drosophila simulans* complex species. Genetics 156:1913-1931.

Klahre, U., A. Gurba, K. Hermann, M. Saxenhofer, E. Bossolini, P. M. Guerin, and C. Kuhlemeier. 2011. Pollinator choice in *Petunia* depends on two major genetic Loci for floral scent production. Curr. Biol. 21:730-739.

Kreitman, M., and M. Aguade. 1986. Genetic uniformity in two populations of *Drosophila melanogaster* as revealed by filter hybridization of four-nucleotide-recognizing restriction enzyme digests. Proc. Natl. Acad. Sci. USA 83:3562-3566.

Kronforst, M. R., L. G. Young, D. D. Kapan, C. McNeely, R. J. O'Neil, and L. E. Gilbert. 2006. Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. Proc. Natl. Acad. Sci. USA 103:6575-6580.

Kulkarni, S. J., A. F. Steinlauf, and J. C. Hall. 1988. The dissonance mutant of courtship song in *Drosophila melanogaster*: isolation, behavior and cytogenetics. Genetics 118:267-285.

Kyriacou, C. P. 2002. Single gene mutations in *Drosophila*: what can they tell us about the evolution of sexual behaviour? Genetica 116:197-203.

Kyriacou, C. P., and J. C. Hall. 1980. Circadian rhythm mutations in *Drosophila melanogaster* affect short-term fluctutaions in the male's courtship song. Proc. Natl. Acad. Sci. USA 77:6729-6733.

Lachaise, D., M. Harry, M. Solignac, F. Lemeunier, V. Benassi *et al.* 2002. Evolutionary novelties in islands: *Drosophila santomea*, a new *melanogaster* sister species from São Tomé. *Proc. R. Soc. Lond. Ser. B* 267:1487–1495.

Lawniczak, M. K., and D. J. Begun. 2004. A genome-wide analysis of courting and mating responses in *Drosophila melanogaster* females. Genome 47:900-910.

Llopart, A., S. Elwyn, D. Lachaise, and J. A. Coyne. 2002. Genetics of a difference in pigmentation between *Drosophil yakuba* and *Drosophila santomea*. Evolution 56:2262-2277.

Lowry, D. B., and J. H. Willis. 2010. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. PLoS ONE 8:e1000500.

Maan, M. E., O. Seehausen, L. Sodderberg, L. Johnson, E. A. Ripmeester, H. D. Mrosso, M. I. Taylor, T. J. van Dooren, and J. J. van Alphen. 2004. Intraspecific sexual selection on s speciation trait, male coloration, in the Lake Victoria cichlid *Pundamilia nyererei*. Proc. Biol. Sci. 271:2445-2452.

Mackay, T. F., S. L. Heinsohn, R. F. Lynman, A. J. Moehring, T. J. Morgan, and S. M. Rollmann. 2005. Genetics and Genomics of *Drosophila* mating behavior. Prac. Natl. Acad. Sci. USA 102:6622-6629.

Marican, C., L. Duportets, S. Birman, and J. M. Jallon. 2004. Female-specific regulation of cuticular hydrocarbons biosynthesis by dopamine in *Drosophila melanogaster*. Insect Biochem. Mol. Biol. 34:823-830.

Matthews, K. A., T. C. Kaufman, and W. M. Gelbart. 2005. Research resources for *Drosophila*: the expanding universe. Nat. Rev. Genet. 6:179-193.

Melo, M. C., C. Salazar, C. D. Jiggins, and M. Linares. 2009. Assortative mating preferences among hybrids offers a route to hybrid speciation. Evolution 63:1660-1665.

Mihola, O., Z. Trachtulec C. Vlcek, J. C. Schimenti, and J. Forejt. 2009. A mouse speciation gene encodes a meiotic histone H3 methyltransferase. Science 323:373-375.

Moehring, A. J., A. Llopart, S. Elwyn, J. A. Coyne, and T. F. Mackay. 2006. The genetic basis of prezygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to mating preference. Genetics 173:215-223.

Moehring, A. J., and T. F. Mackay. 2004. The quantitative genetic basis of male mating behavior in *Drosophila melanogaster*. Genetics 167:1249-1263.

Moehring, A. J., J. Li, M. D. Schug, S. G. Smith, M. deAngelis, T. F. Mackay, and J. A. Coyne. 2004. Quantitative trait loci for sexual isolation between *Drosophila simulans* and *D. mauritiana*. Genetics 167:1265-1274.

Morehouse, N. I., and R. L. Rutowski. 2010. In the eyes of the beholders: female choice and avian predation risk associated with an exaggerated male butterfly color. Am. Nat. 176:768-784.

Moulin, B., T. Aubin, and J. M. Jallon. 2004. Why there is a one-way crossability between *D. melanogaster* and *D. simulans*? An ontogenic explanation. Genetica 120:285-292.

Nanda, P., and B. N. Singh. 2011. Effect of chromosome arrangements on mate recognition system leading to behavioral isolation in *Drosophila ananassae*. Genetica 139:273-279.

Narraway, C., J. Hunt, N. Wedell, and D. J. Hosken. 2010. Genotype-by-environment interactions for female preference. J. Evol. Biol. 23:2550-2557.

Neckameyer, W. 1998. Dopamine modulates female sexual receptiveness in *Drosophila melanogaster*. J. Neurogenet. 12:101-114.

Nickel, D., and A. Civetta. 2009. An X chromosome effect responsible for asymmetric reproductive isolation between male *Drosophila virilis* and heterospecific females. Genome 52:49-56.

Noor, M. A. 1995. Speciation driven by natural selection in *Drosophila*. Nature 375:674-675.

Noor, M. A., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. Prac. Natl. Acad. Sci. USA 98:12084-12088.

O'Dell, K. M., J. D. Armstrong, M. Y. Yang, and K. Kaiser. 1995. Functional dissection of the *Drosophila* mushroom bodies by selective feminization of genetically defined subcompartments. Neuron 15:55-61.

Ortiz-Barrientos, D., B. A. Counterman, and M. A. Noor. 2004. The genetics of speciation by reinforcemnt. PLoS Biol. 2:e416.

Ortiz-Barrientos, D., B. A., and M. A. Noor. 2005. Evidence for a one-allele assortative mating locus. Science 310:1467.

Peixoto, A. A., and J. C. Hall. 1998. Analysis of termperature-sensitive mutants reveals new genes involved in the courtship of *Drosophila*. Genetics 148:827-838.

Pineiro, R., M. C. Carracedo, J. I. Izquierdo, and P. Casares. 1993. Bidirectional selection for female receptivity in *Drosophila melanogaster*. Behav. Genet. 23:77-83.

Price, C. S. 1997. Conspecific sperm precedence in *Drosophila*. Nature 388:663-666.

Ritchie, M. G., and C. P. Kyriacou. 1994. Reproductive isolation and the period gene of *Drosophila*. Mol. Ecol. 3:595-599.

Ritchie, M. G., and M. A. Noor. 2004. Evolutionary genetics: gene replacement and the genetics of speciation. Heredity 93:1-2.

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 evoltion in *D. melanogaster* and *D. simulans*. Genetica 120: 195-212.

Ruedi, E. A., and K. A. Hughes. 2009. Age, but not experience, affects courtship gene expression in male *Drosophila melanogaster*. PLoS one.

Rundle, H. D., S. F. Chenoweth, P. Doughty, and M. W. Blows. 2005. Divergent selection and the evolution of signal traits and mating preferences. PLoS Biol. e368.

Sakai, T., and N. Ishida. 2001. Time, love, and species. Neuro. Endocrinol. Lett. 22:222-228.

Sawamura, K., H. Zhi, K. Setoguchi, H. Yamada, T. Miyo, M. Matsuda, and Y. Oguma. 2008. Genetic analysis of female mating recognition between *Drosophila ananassae* and *Drosophila pallidosa*: application of interspecific mosaic genome lines. Genetica 133:179-185.

Servedio, M. R., and M. A. F. Noor. 2003. The role of reinforcement in speciation: theory and data. Ann. Rev. Ecol. Evol. 34:339-364.

Stevison, L. S., K. B. Hoehn, and M. A. Noor. 2011. Effects of inversions on within- and between-species recombination and divergence. Genome Biol. Evol. 3:830-841.

Suzuki, K., N. Juni, and D. Yamamoto. 1997. Enhanced mate refusal in female *Drosophila* induced by a mutation in the *spinster* locus. Appl. Entomol. Zool. 32:235-243.

Ting, C. T, A. Takahashi, and C. I. Wu. 2001. Incipient speciation by sexual isolation in *Drosophila*: concurrent evolution at multiple loci. Proc. Natl. Acad. USA 98:6709-6713.

Tomaru, M., H. Yamada, and Y. Oguma. 2004. Female mate recogniction and sexual isolation depending on courtship song in *Drosophila sechellia* and its sibling species. Genes Genet. Syst. 79:145-150.

Tsacas, L., and J. David. 1974. *Drosophila mauritiana* n. sp. du groupe *melanogaster* de l'ile Maurice. Bull. Soc. Entomol. Fr. 79:42-46.

Uenoyama, T., and Y. Inoue. 1995. Genetic studies on premating isolation in *Drosophila simulans*. I.A *D. simulans* line highly crossable with *D. melanogaster*. Jpn. J. Genet. 70:365-371.

Weetman, D., C. S. Wilding, K. Steen, J. Pinto, and M. J. Donnelly. 2011. Gene flow-dependent genomic divergence between *Anopheles gambiae* M and S forms. Mol. Biol. Evol. Epub ahead of print.

Wheeler, D. A., C. P. Kyriacou, M. L. Greenacre, Q. Yu, J. E. Rutila, M. Rosbash, and J. C. Hall. 1991. Molecular transfer of a species-specific behaviour from *Drosophila simulans* to *Drosophila melanogaster*. Science 251:1082-1085.

Wicker-Thomas, C., and J. M. Jallon. 2001. Control of female pheromones in *Drosophila melanogaster*

Wood, D., J. M. Ringo, and L. L. Johnson. 1980. Analysis of courtship sequences of the hybrids between *Drosophila melanogaster* and *Drosophila simulans*. Behav. Genet. 10:459-466.

Wu, C. I., H. Hollocher, D. J. Begun, C. F. Aquardro, Y. Xu, and M. L. Wu. 1995. Sexual isolation in *Drosophila melanogaster*: a possible case of incipient speciation. Proc. Natl. Acad. USA 92:2519:2523.

Yamamoto D., and Y. Nakano. 1999. Sexual behavior mutants revisited: molecular and cellular basis of *Drosophila* mating. Cell. Mol. Life Sci. 56:634-646.

Chapter 2: Mapping of the right arm of the third chromosome for genes for interspecific female preference in *D. simulans*.

Introduction

The process of species divergence creates distinct evolutionary lineages, giving rise to biodiversity. Although this process of divergence is not fully understood, species divergence depends upon reproductive barriers that reduce or eliminate gene flow between two populations that otherwise have the potential to interbreed. Of these barriers, prezygotic isolation (e.g. behavioural isolation) is considered to be the first to evolve, and is therefore the primary component leading to species divergence (Coyne and Orr 1997). The identification of the genetic architecture of behavioural isolation is, therefore, a critical component necessary for our understanding of the process of speciation, yet to date no individual genes for behavioural isolation have been identified.

Divergent mating behaviours between closely-related species have been shown to isolate interbreeding population of *Drosophila* (Noor *et al.* 2001), butterflies (Kronforst *et al.* 2006), and even plants (Klahre *et al.* 2011), but is typically very difficult to investigate because in animals male courtship signaling and female receptivity, like most quantitative traits, have multiple sources of variation ranging from genetic to environmental (Narraway *et al.* 2010). Many of the difficulties involved in studying behavioural isolation can be bypassed by using the *Drosophila* model system. The mating behaviour of *Drosophila* has been well investigated, involves many stereotypical and sequential phenotypes, and contains

multimodal signaling (Hall 1994). Courtship behaviour progresses from the orientation of the male to the female, the male tapping the female's abdomen with his front leg (exchange of chemical signals), the male vibrating his wing to produce a species-specific courtship song, licking the female's genitalia (exchange of chemical signals), and finally attempting copulation by curling his abdomen towards the female. If the male is successful at initiating copulation, the female will slow locomotion, orient to the male, present her abdomen and allow for copulation (Hall 1994; Greenspan 1995; Greenspan and Ferveur 2000). In addition to the well-characterized courtship and mating behaviour of *Drosophila*, this species group also has the advantage of powerful genetic tools. Twelve species of *Drosophila* have had their genome fully sequenced (*Drosophila* 12 genomes consortium 2007), and one species (*D. melanogaster*) has multiple gene mutant lines available that can be used to facilitate the identification of individual genes contributing to behaviour (Moehring and Mackay 2004; Mackay *et al.* 2005; Carney 2007; Ruedi and Hughes 2008). Lastly, a large number of individuals can be kept in a controlled and uniform environment in order to reduce environmental variance.

Mutagenesis studies, with use of the *Drosophila* model organism, have contributed greatly to the knowledge of the genetic basis of mating behaviours. Many genes for male mating behaviour have been identified and have provided insight into the sensory systems used during courtship such as female pheromone detection (*voila*, Balakireva *et al.* 1998) male pheromone production (*nerd*, Ferveur 2005), and male courtship song production (*period*, Kyriacou and Hall 1980; *cacophony*, Kulkarni *et al.* 1988; *slowpoke*, *paralytic*, and *no action potential*, Peixoto

and Hall 1998). Three genes for female receptivity have been identified (*spinster,* Suzuki *et al.* 1997; *dissatisfaction*, Finley *et al.* 1997; *chaste*, Juni and Yamamoto 2009) and all three have been shown to be expressed in the brain. However, genes identified through these studies do not seem to contribute to interspecific mating behaviour (Moehring *et al.* 2004, 2006).

Although no genes have been identified for behavioural isolation in *Drosophila*, genomic regions have been identified that contribute to the trait, and two main patterns have emerged. First, the third chromosome (or homologous region to the *melanogaster* subgroup) has a disproportionately large contribution to behavioural isolation (Coyne 1992; Uenoyama and Inoue 1995; Carracedo *et al.* 1995; Doi *et al.* 2001; Noor *et al.* 2001; Ting *et al.* 2001; Moehring *et al.* 2004; Moehring *et al.* 2006; Sawamura *et al.* 2008), with a greater contribution of the right than the left arm (Noor *et al.* 2001; Doi *et al.* 2001; Moehring *et al.* 2004, 2006; Sawamura *et al.* 2008). Second, genes contributing to behavioural isolation are more likely to be found in regions of low recombination, such as those near the centromeres or telomeres and those found in interspecific inversion polymorphisms (Noor *et al.* 2001; Nanda and Singh 2001; Sawamura *et al.* 2008; Sousa-Neves and Rosas 2010). However, without individual genetic loci it is difficult to characterize the true underlying basis of behavioural isolation.

One reason that progress has been limited is that the species of *D. melanogaster* is the system with the vast majority of the genetic tools, and is the only *Drosophila* system where individual genes can easily be tested with pre-existing mutant lines. Unfortunately, this species produces sterile offspring with its sibling

species (Sturtevant 1920), thus preventing the use of recombination mapping to identify candidate regions. This type of mapping requires the crossing of fertile F_1 hybrids to one of the parental species. In the species pairs where backcrossing is possible (such as *D. pseudoobscura* and *D. persimilis*, *D. ananassae* and *D. pallidosa*, and *D. simulans* and *D. mauritiana*) genomic regions have been identified for behavioural isolation such as an entire chromosome (Coyne 1989, 1992), a specific arm of a chromosome (Noor *et al.* 2001; Sawamura *et al.* 2008), sub-chromosomal regions (Doi *et al.* 2001; Ting *et al.* 2001), and smaller genomic regions identified through QTL studies (Moehring *et al.* 2004, 2006) in these other *Drosophila* species pairs. However, refinement of these regions is dependent upon the limited ability of recombination mapping to reduce regions to a very small number of individual genetic loci, since no readily-available mutants of individual genes exist in these other *Drosophila* species.

Here, I circumvented this problem by using *D. melanogaster*'s unique tools to identify refined genomic regions underlying *D. simulans* female preference against *D. melanogaster* males, and the methodology I used can be expanded to subsequently fine map these regions to individual genetic loci. *D. melanogaster* and its sister species *D. simulans* diverged approximately 5.4 million years ago (Tamura *et al.* 2004). These two species differ by roughly 3% of their genetic sequence and a large interspecific inversion polymorphism on the right arm of the third chromosome (Ranz *et al.* 2007). The behavioural isolation between the two species is asymmetrical: female *D. melanogaster* will readily mate with *D. simulans* males but the converse pairing very rarely occurs (Moulin *et al.* 2004). Since *D. melanogaster*

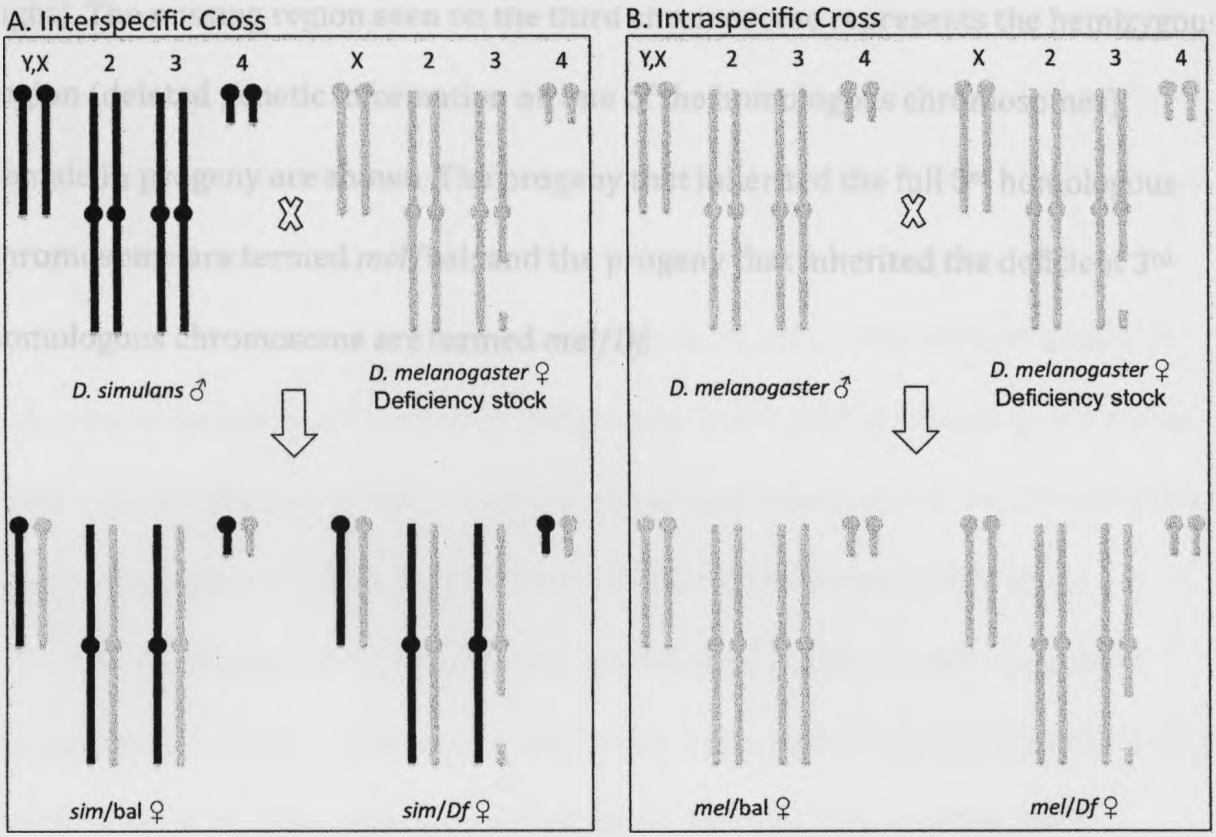
males fully court *D. simulans* females, the behavioural isolation between this species pair is primarily due to the female's rejection of these courting heterospecific males (Jamart *et al.* 1993). Crosses in the permissive direction, between *D. melanogaster* females and *D. simulans* males, produce only sterile F_1 females (Sturtevant 1920).

Although the F_1 female hybrids are sterile, they are fully viable, and their behaviour can be assessed. Since the loci for *melanogaster*-like receptivity are dominant over those for *simulans*-like rejection (Davis *et al.* 1996), these F_1 females are receptive to *D. melanogaster* male courtship. This permits the use of a *D. melanogaster* genetic tool that is widely used to locate recessive genes with large effect sizes and can effectively map a trait down to single gene(s): deficiency mapping.

The process of deficiency mapping requires the use of deficiency lines. These lines are almost entirely diploid except for a small portion that is hemizygous (having only one copy of the genes) due to a deletion ('deficiency') on one of the homologous chromosomes. This deficiency is maintained over a 'balancer' chromosome that has multiple inversions that prevent the recovery of offspring that have had recombination on that chromosome as well as a dominant visible mutation that is homozygous lethal. Each line has a different hemizygous region, and these deficiencies exist for almost all regions of the genome.

When *D. simulans* are crossed with a *D. melanogaster* deficiency lines (Figure 2.1), the F_1 female hybrids that are produced are heterozygous (heterospecific) throughout their genome except for the region with the deletion. Here, the only alleles that can be expressed will originate from *D. simulans*, exposing a small region

Figure 2.1: Inter- and intraspecific crosses with *D. melanogaster* deficiency lines.



Crosses used in deficiency mapping for interspecific female preference.

A) Interspecific cross. Interspecific hybrids are created to compare the mating behaviour of the two hybrids. Wild *D. simulans* males (shown in black) are crossed with *D. melanogaster* female from a deficiency line (shown in grey). The missing region seen on the third chromosome represents the hemizygous region (deleted genetic information on one of the homologous chromosomes). Female F_1 progeny are shown. The progeny that inherited the full 3rd homologous chromosome are termed *sim*/bal; and the progeny that inherited the deficient 3rd homologous chromosome are termed *sim/Df.* B) Intraspecific cross. Full species hybrids are created to control for the effect of the balancer and deficiency chromosome on mating behaviour. Wild *D. melanogaster* males (shown in grey on the left) are

crossed with *D. melanogaster* female from a deficiency line (shown in grey on the right). The missing region seen on the third chromosome represents the hemizygous region (deleted genetic information on one of the homologous chromosomes). Female F_1 progeny are shown. The progeny that inherited the full 3rd homologous chromosome are termed *mel*/bal; and the progeny that inherited the deficient 3rd homologous chromosome are termed *mel*/*Df.*

of *D. simulans* genome in an otherwise heterozygous background. The dominance of *melanogaster* alleles means that F_1 females should be receptive to *D. melanogaster* male courtship. This is also true in F_1 females that inherit a deficiency, unless the exposed region (harbouring only *D. simulans* alleles) contains genes contributing to *simulans*-like rejection, in which case the females should not copulate with a *D. melanogaster* male. These regions can, therefore, be used to map where genes for behavioural isolation are located in the genome. Individual candidate genes within these regions can then be tested using the same premise as above, but by using lines with a single gene disrupted rather than a small region with multiple genes.

Deficiency mapping has previously been used in this manner to uncover genes associated with a variety of quantitative traits, including viability and fertility (Presgraves *et al.* 2004; Sawamura *et al.* 2004), sterility (Perotti *et al.* 2001), longevity (Pasyukova *et al.* 2000), mating behaviour (Moehring and Mackay 2004), and myoblast fusion (Bour *et al.* 2000). Here, I used deficiency mapping to identify small genomic regions that contribute to the behavioural isolation of *D. simulans* females against heterospecific *D. melanogaster* males.

I focused my efforts on the right arm of the third chromosome due to large effect of this region on behavioural isolation between *Drosophila* species (Doi *et al.* 2001; Noor *et al.* 2001; Moehring *et al.* 2006; Sawamura *et al.* 2008). I also tested the regions harbouring two genes previously identified through mutagenesis as affecting female receptivity within *D. melanogaster*: *spinster* (Suzuki *et al.* 1997), and *dissatisfaction* (Finley *et al.* 1997). The method of deficiency mapping allows for the identification of candidate genes for both female preference and behavioural

isolation between species, as well as a refined assessment of how these genes are distributed on the chromosome.

Materials and Methods

Stocks

Wild-type stocks of *D. simulans* (Florida City) were obtained from the Drosophila Species Stock Center (San Diego, CA, USA); wild-type *D. melanogaster* (BJS1) were collected in 2009 in London, ON, Canada by Dr. Brent Sinclair. *D. melanogaster* stocks with deficiencies spanning the right arm of the third chromosome (Figure 2.1), as well as the two regions containing the genes for female receptivity that had previously been identified through mutagenesis (*spinster*, Suzuki *et al.* 1997; *dissatisfaction*, Finley *et al.* 1997), were obtained from the Bloomington Drosophila Stock Center (Bloomington, IN, USA; Table 2.1). In these deficiency stocks, one homolog of the third chromosome has a deficiency *(Df)* and one homolog is the balancer (bal). All of the deficiency breakpoints were provided by the Bloomington Drosophila Stock Center. All stocks and crosses were maintained in incubators with a 14:10 hour light:dark cycle, 25°C, and approximately 80% relative humidity. All stocks and crosses were maintained in standard 8-dram food vials containing approximately 7ml of the Bloomington Drosophila Stock Center's standard cornmeal/yeast media recipe. For three lines that did not survive well at 25°C (*Df*(3R),crb87-5, *Df*(3R)e-R1, and *Df*(3R)ry85), crosses were maintained as above, but at 21°C.

Table 2.1: Deficiency lines used to map 3R and test previously identified female preference genes.

Lines that spanned 3R

Deficiency	Region	Balancer	Marker	Notes
Df(3R)ME15	81F3-6;82F5-7	MKRS	Sb	completed
Df(3R)ED5156	82F8;83A4	TM6c	Sb	completed
Df(3R)Exel6144	83A6;B6	TM6b	Tb	completed
Df(3R)ED5177	83B4;B6	TM6c	Sb	completed
Df(3R)BSC464	83B7;E1	TM6c	Sb	completed
Df(3R)Tpl10	83C1-2;84B1-2	TM3	Sb	completed
Df(3R)Antp17	84A5;D9	TM3	Sb, Ser	too weak
Df(3R)p712	84D4-6;85B6	TM3	Sb, Ser	too weak
Df(3R)dsx37	84D8;85B3-5	TM3	Sb, Ser	too weak
Df(3R)ED5330	85A5;D1	TM6c	Sb	completed
Df(3R)BSC666	85C2;D11	TM6c	Sb	inviable
Df(3R)by62	85D10-11;F8-11	TM1	Me	too weak
Df(3R)GB104	85D12;E10	TM3	Sb, Ser	inviable
Df(3R)BSC38	85F1-2;86C7-8	TM2	Ubx	completed
Df(3R)M-Kx1	86C1;87B1-5	TM3	Sb	completed
Df(3R)T-32	86E2-4;87C6-7	MRS	Sb	completed
Df(3R)ry85	87B15-C;F15-88	MKRS	Sb	21 incubator
Df(3R)ED5612	87C7;F6	TM6c	Sb	inviable
Df(3R)ED5644	88A4;C9	TM6c	Sb	48-hour
Df(3R)red1	88B1;D3-4	TM1	Me	too weak
Df(3R)ED5664	88D1;E3	TM6c	Sb	completed
Df(3R)BSC471	88E3;E5	TM6c	Sb	completed
Df(3R)BSC741	88E8;F1	TM6c	Sb	too weak
Df(3R)ED5705	88E12;89A5	TM2	Ubx	too weak
Df(3R)sbd105	88F9-A;89B9-10	TM3	Ser	completed
Df(3R)P115	89B7-8;E7	TM1	Me	completed
Df(3R)RD31	89E2;90D	In(3R)C	Sb	completed
Df(3R)DG2	89E1-F4;91B1-2	TM2	Ubx	completed
Df(3R)ED5780	89E11;90C1	TM2	Ubx	completed
Df(3R)Cha7	90F1-4;91F5	TM6b	Tb	completed
Df(3R)Dl-BX12	91F1-2;92D3-6	TM6b	Tb	completed
Df(3R)H-B79	92B3;F13	TM2	Ubx	completed
Df(3R)BSC43	92F7-93A1;B3-6	TM2	Ubx	completed
Df(3R)e-N19	93B;94	TM2	Ubx	too weak
Df(3R)e-R1	93B6-7;93D2	TM3	Sb, Ser	21 incubator
Df(3R)e-GC3	93C6;94A4	TM6b	Tb	48-hour
Df(3R)hh	93F11;94D13	Sb	Sb	too weak
Df(3R)ED6096	94B5;E7	TM6c	Sb	completed
Df(3R)BSC56	94E1-2;F1-2	TM2	Ubx	completed

Deficiency	Region	Balancer	Marker	Notes
Df(3R)BSC137	94F1;95A4	TM6b	Tb	completed
Df(3R)BSC489	94F3;95D1	TM6c	Sb	completed
Df(3R)Exel6196	95C12;D8	TM6b	Tb	completed
Df(3R)crb-F89-4	95D7-11;F15	TM3	Sb	inviable
Df(3R)Exel6197	95D8;E1	TM6b	Tb	completed
Df(3R)ED6187	95D10;96A7	TM2	Ubx	completed
Df(3R)crb87-5	95F7;96A17-18	TM3	Ser	21 incubator
Df(3R)ED6220	96A7;C3	TM6c	Sb	completed
Df(3R)BSC461	96B15;D1	TM6c	Sb	completed
Df(3R)Exel6202	96D1;D1	TM6b	Tb	completed
Df(3R)Exel6203	96E2;E6	TM6b	Tb	completed
Df(3R)BSC321	96E6;E9	TM6c	Sb	completed
Df(3R)BSC140	96F1;F10	TM6b	Tb	completed
Df(3R)Espl3	96F1;97B1	TM6c	Sb, Tb	completed
Df(3R)Tl-P	97A;98A1-2	TM3	Ser	inviable
Df(3R)BSC497	97E6;98B5	TM6c	Sb	completed
Df(3R)IR16	97F1-2;98A	TM3	Sb	completed
Df(3R)BSC567	98B6;E5	TM6c	Sb	inviable
Df(3R)3450	98E3;99A6-8	TM6b	Tb	completed
Df(3R)ED6316	99A5;C1	TM6c	Sb	inviable
Df(3R)BSC547	99B5;C2	TM6c	Sb	completed
Df(3R)L127	99B5-6;F1	TM6	Ubx	completed
Df(3R)BSC620	99C5;D3	TM6c	Sb	completed
Df(3R)B81	99D3;3Rt	TM3	Sb	completed
Lines that tested previously identified female receptivity genes				
dissatisfaction				
Df(2L)cl-h3	25D2-4;26B2-5	SM6b	Cy	completed
PBac{w[+mC]=WH}	26A1;26A2	CyO	Cy	completed
spinster				
Df(2R)Jp1	51D3-8;52F5-9	CyO	Cy	completed
P{ry[+t7.2]=neoFRT}43D	52E6;52E7	CyO	Cy	completed

Completed = line was successfully completed; Inviable = *sim*/bal and/or *sim*/*Df* was inviable; Too weak = line was too weak to be maintained; 21 incubator = crosses were completed within a 21°C incubator; 48-hour = hybrids were paired with males for 48 hours before dissections.

Crosses

Virgin females from the deficiency *D. melanogaster* stocks were stored in groups of 1-20 and aged for 7-18 days. Virgin males of wild-type *D. simulans* and *D. melanogaster* were stored in groups of 1-20 and aged for 0-7 days. To create interspecies hybrids, ten aged virgin females were paired with 20-25 wild-type aged *D. simulans* males; the volume of the vial was reduced to increase contact between the males and females. This created two types of interspecies hybrid offspring: *sim/Df* and *sim*/bal. To create pure-species control hybrids, five aged virgin females were paired with five wild-type aged *D. melanogaster* males. This created two types of intraspecies offspring: *mel/Df* and *mel*/bal. Crosses were transferred to new food vials every 5 days until no more larvae were produced.

Mating assay

Hybrid and pure species female offspring from the inter- and intraspecific crosses were collected 0-8 hours after eclosion to ensure virginity, separated using the dominant balancer marker phenotype under brief CO_2 anaesthesia, and transferred to new vials at low densities of 1-10 flies. One 5-7 day old virgin female was placed with one 5-7 day old virgin wild-type *D. melanogaster* male for 45 minutes in an 8 dram glass vial that had been misted with water to provide humidity. An equal number of the four genotypes (*sim/Df, sim*/bal, *mel/Df, mel*/bal) were assayed at the same time in order to control for environmental effects. The time that courtship was initiated (courtship occurrence) and that copulation was initiated (copulation occurrence) were recorded. From these measures, the proportion of the total number of females that were courted, the proportion that

copulated only out of those that were courted, and the courtship latency was calculated.

Although interspecies hybrids will mate with *D. melanogaster* males, this is at a reduced rate, and very few interspecies matings occurred (for both *sim/Df* and *sim*/bal) within the 45 minute observation period for some of the lines that were tested. To get a more thorough measure of the differences between mating values for *sim/Df* compared to *sim*/bal, I also placed the interspecies hybrid female and her paired *D. melanogaster* male in a vial containing food after completion of the behaviour assay for 30 of the lines. After 24-28 hours, the female reproductive tract and spermathecae were dissected and scored for sperm presence under a light microscope. For two lines with very low interspecies copulation occurrence for both *sim/Df* and *sim*/bal (*Df*((3R)e-GC3 and *Df*(3R)ED5644), the interspecies hybrid female and their paired *D. melanogaster* male were placed in a vial containing food after completion of the behaviour assay and held for 48-50 hours, at which time the female reproductive tract and spermathecae were dissected and scored for sperm presence under a light microscope.

Data Analysis

For each line, a G test ($\alpha<0.05$) was used for the binomial variables of courtship occurrence. However, if the minimum expected value for any category was lower than five in the G test, then a Fishers Exact test was used ($\alpha<0.05$). Also, an ANOVA ($\alpha<0.05$) was used for the continuous variables of courtship latency. Both of these were used to test for the potential effect of male courtship behaviour that could potentially bias the interpretation of the female mating behaviour. If all

females were not courted equally, the difference could be misinterpreted as a reluctance to mate on the female's behalf.

The balancer chromosome in the deficiency lines has multiple inversion breakpoints that could potentially disrupt the normal functioning of the *D. melanogaster* gene. If this occurred, this disruption would act as another deficiency uncovering the functioning *D. simulans* allele. If this allele functioned as a gene for interspecific female preference, these *sim*/bal females would be less likely to mate with *D. melanogaster* males compared to the *sim/Df* females of the same line. To test for this possibility, I grouped the *sim*/bal females of the different lines according to their balancer and compared the copulation occurrence (copulated or did not copulate) with the *D. melanogaster* male for the 45 minute assay and the 24 hour assay with a one-way ANOVA ($\alpha<0.05$). This was only possible for balancers where more than one line with the balancer was tested. In order to assess if the balancer's effect on mating was due to the breakpoints and not the genetic background of the lines, I grouped the *sim/Df* females of the different lines according to their balancer and compared the copulation occurrence (copulated or did not copulate) with the *D. melanogaster* male for the 45 minute assay and the 24 hour assay with a one-way ANOVA ($\alpha<0.05$).

To test the different regions of the right arm of the third chromosome as well as the two previously identified genes for female receptivity, a G test ($\alpha<0.05$) was used for the binomial variables of copulation occurrence. However, if the minimum expected value for any category was lower than five in the G test, then a Fisher's Exact test was used ($\alpha<0.05$). For a result to be significant, the *sim/Df* hybrids must

show a reduced amount of mating compared to the interspecies control (*sim*/bal) after those values are corrected for any effect of the deficiency and balancer genetic background (*mel/Df, mel*/bal). As my sample sizes were of moderate size, and behaviour is thought to be affected by multiple genes of small effect (Greenspan 1995), I also considered a line to be 'potentially significant' if $0.15>\alpha>0.05$ and in the expected direction.

When a region of the genome was located that caused a significant effect on the behaviour, further refinement was completed by interline comparison. This is possible due to the large number of available deficiency lines, and that some regions uncovered by deficiencies are also partially uncovered by other deficiencies. Thus, if deficiency line A has no significant difference in behaviour and deficiency line B does, the portion of the deficiency of A that overlaps with B must not house any genes for female preference. This decreases the size of the region considered to most likely contains genes for interspecific female preference. Likewise, if two overlapping lines are both significant, it is most likely the region of overlap that contains the genes contributing to the trait. Details about the location, function, and biological properties of the candidate genes within the refined regions were identified using the FlyBase online database (Tweedie *et al.* 2009; www.flybase.org).

Results

I used deficiency mapping to locate genes for interspecific female preference in *D.* simulans. I compared the mating behaviour of four types of females (*sim*/bal, *sim/Df, mel*/bal, and *mel/Df*) when paired with *D. melanogaster* males. If the *sim/Df*

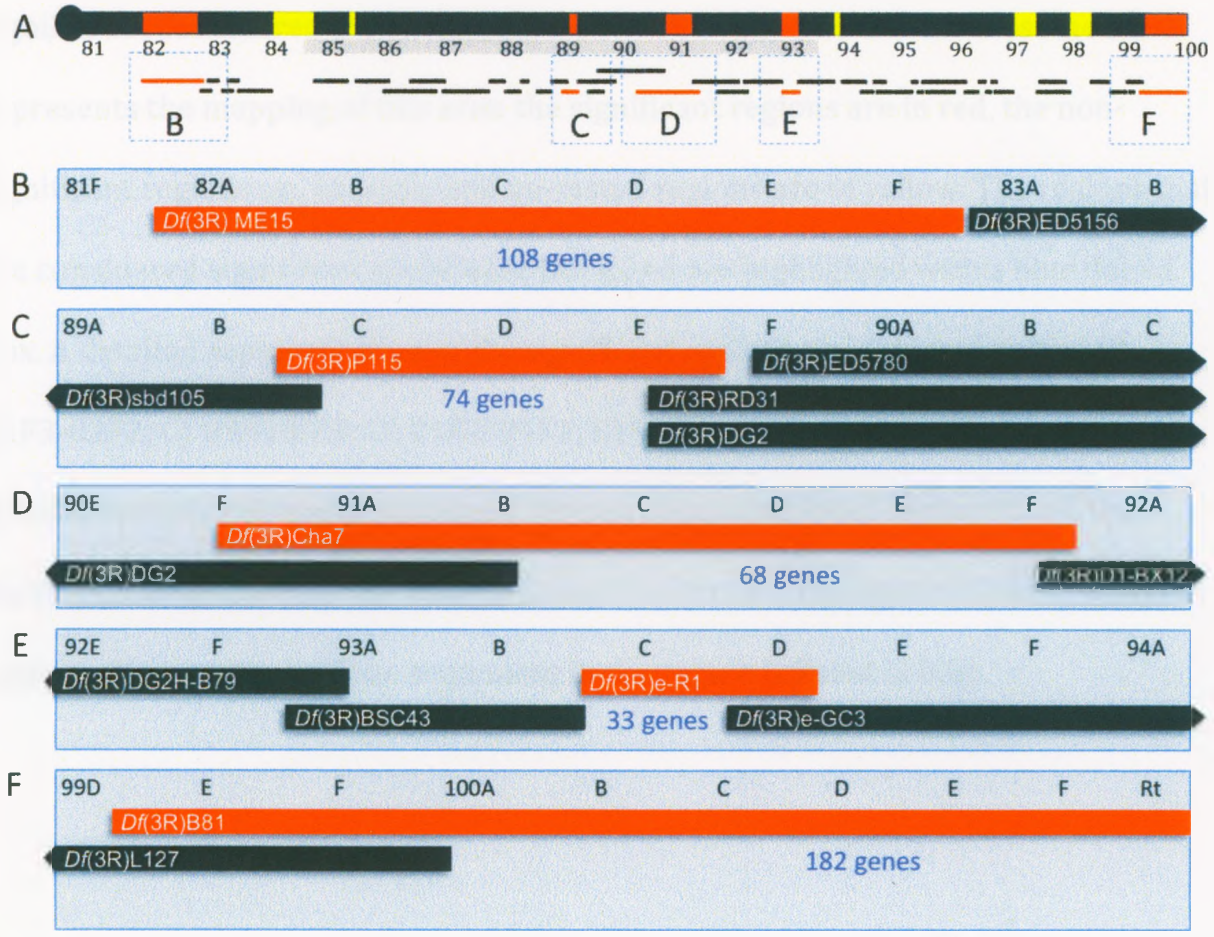
females' mating propensity did not significantly differ from that of the *sim*/bal or if this difference was also observed in the control females, then the null hypothesis was not rejected and, therefore, the deficiency did not uncover a gene(s) for interspecific female preference. However, if the *sim*/*Df* females mated significantly less than the *sim*/bal and this difference is significantly different than that of the control females, the null hypothesis was rejected. This would indicate that the deficiency uncovered a gene(s) for behavioural isolation.

I attempted to complete 67 lines. However, nine lines were not completed because the interspecies hybrids were too weak for behaviour assays, and an additional seven lines could not be completed because either *sim*/bal or *sim*/*Df* or both were inviable (Table 2.1). The 47 lines I successfully crossed covered 80.3% of genes and 82.6% of the genome on the right arm of the third chromosome (Figure 2.2), with four additional lines tested that overlapped two previously identified female preference genes. The untested cytological regions are 83A5, 84B3-85A4, 85D2-E15, 88C10-11, 88E6-F9, 94A5-B4, 97B2-E5, 98B6-E2, and 99A7-B3. These regions include 699 genes.

Mating

For each line, I compared the copulation occurrence (proportion that copulated) of a *D. melanogaster* male when paired with a female of one of the four genotypes for the 45 minute observational assay and the 24 hour sperm assay using either a G test or a Fisher's Exact test ($\alpha<0.05$, or 'potentially significant' if $0.15>\alpha>0.05$). When I tested the two genes previously identified as affecting female intraspecific mate preference, I did not find a significant effect for interspecific mate

Figure 2.2: Overlapping deficiencies spanning 3R used to locate genes for interspecific female preference in *D. simulans*.



A representation of the deficiencies used to map interspecific mating preference in *D. simulans* females. A) The bar represents the right arm of the third chromosome. To the left, the circle represents the centromere and the blunt right end represents the telomere. The numbers under the bar are the approximate cytological regions and the gray bar represents the interspecific inversion polymorphism. Under the chromosome the deficiencies that covered 3R are represented as bars: the non-significant lines are black and the significant lines are red. The interline comparisons were used to decrease genomic regions that were considered significant. Some of the lines have specific breakpoints listed (base pairs) and others have estimated

breakpoints listed (cytological regions). In order to include all possible candidate genes, when a line was significant I considered the largest area and when it was not significant I considered the smallest area. The coloration of the chromosome bar represents the mapping of this arm: the significant regions are in red, the non-significant regions are in black, and un-tested regions are in yellow. The regions that are considered significant are labeled Bto E and are highlighted with a blue dahed box. A detailed representation of the significant regions are depicted below: B) 81F3-82F2; C) 89B10-E2; D) 91B2-91F1; E) 93B7-93C5; F) 99F2-3Rt. The deficiencies are represented as bars: the non-significant lines are black and the significant lines are red. The numbers over the bars are the approximate cytological regions. The number of gene sequences in the region is listed in blue.

preference during the 45 minute behaviour assay or after the 24 hour sperm assay (Table 2.2). For the deficiency lines spanning the right arm of the third chromosome, no lines were found to be significant in the expected direction for copulation occurrence with the 45 minute assay (Table 2.2), but four lines were considered potentially significant: *Df*(3R)M-Kx1 (Fisher's Exact test p=0.14), *Df*(3R)P115 (Fishers Exact test p=0.11), *Df*(3R)e-R1 (Fisher's Exact test p=0.12) and *Df*(3R)B81 (G=3.03, p=0.08). Four lines were found to be significant with the 24 hour sperm assay: *Df*(3R)ME15 (G=5.93, p=0.01), *Df*(3R)P115 (G=11.32, p=0.0008), *Df*(3R)Cha7 (Fisher's Exact test p=0.03), and *Df*(3R)e-R1 (Fisher's Exact test p=0.03). An additional line, *Df*(3R)B81, was not strictly significant (G=3.34, p=0.067);since this line also approached significance with the behaviour assay, even with my relatively small sample sizes, hereafter I will consider this line as significant, with an acknowledgement for the need to confirm with additional tests and other overlapping deficiencies.

Male Courtship Behaviour

To ensure that the differences in copulation were due to female preference and not due to a reduction in male courtship of those females, I also compared courtship occurrence and courtship latency for the lines. Approximately 82% of the females were courted within the 45 minute assay period (*sim/Df*=83%, *sim*/bal=79%, *mel/Df*=85%, *mel*/bal=80%). For each line, I compared the courtship occurrence (courted or not courted) by the *D. melanogaster* male towards the female of each of the four genotypes (*sim*/bal, *sim/Df, mel*/bal, and *mel/Df)* with use of a G test or a Fishers Exact test ($p<0.05$). No lines produced a significant value

Table 2.2: Mating occurrence of the four types of females with *D. melanogaster* males.

Lines that spanned 3R

		mel/bal		*mel*/*Df*		*sim*/bal		*sim*/*Df*			Sperm Assay			
***Df*(3R)**	**N**	**Crt**	**Cop**	**Crt**	**Cop**	**Crt**	**Cop**	**Crt**	**Cop**	**G test/Fishers**	**N**	**bal**	***Df***	**G test/Fisher's**
ME15	20	17	12	18	15	17	1	13	0	G=1.79 p=0.18 p=0.48	20	13	3	G=5.93 p=0.02
ED5156	22	20	14	20	10	21	2	19	1	G=0.04 p=0.84 p=1	20	7	6	G=0.07, p=0.80
Exel6144	20	17	13	17	14	20	0	15	0	G=0 p=1 p=1	0	NA	NA	
ED5177	20	19	13	17	13	17	1	16	1	G=0.01 p=0.97 p=1	0	NA	NA	
BSC464	20	18	10	18	14	14	0	19	0	G=0 p=1 p=1	0	NA	NA	
Tpl10	7	4	3	6	6	5	3	7	6	G=0.01 p=0.94 p=1	0	NA	NA	
ED5330	20	16	11	17	13	15	0	18	0	G=0 p=1 p=1	20	7	9	G=0.02 p=0.90
BSC38	20	17	13	18	12	16	1	16	3	G=1.43 p=0.23 p=0.33	0	NA	NA	
M-Kx1	20	12	9	16	9	17	5	18	0	G=5.76 p=0.02 p=0.138	32	16	11	G=0.38 p=0.54
T-32	20	13	9	14	10	20	4	14	1	G=0.06 p=0.80 p=1	0	NA	NA	
ry85	20	16	9	16	12	16	2	17	4	G=0.16 p=0.69 p=0.68	20	13	13	G=0.24 p=0.63
ED5644	28	25	16	26	22	27	3	26	1	G=1.52 p=0.22 p=0.33	36	18	17	G=0.71 p=0.40

		mel/bal		*mel*/*Df*		*sim*/bal		*sim*/*Df*			Sperm Assay			
***Df*(3R)**	**N**	**Crt**	**Cop**	**Crt**	**Cop**	**Crt**	**Cop**	**Crt**	**Cop**	**G test/Fishers**	**N**	**bal**	***Df***	**G test/Fisher's**
ED5664	20	13	6	16	12	11	0	16	2	G=2.25 p=0.13 p=0.53	20	3	10	G=0.13 p=0.72 p=1
BSC471	20	17	13	15	10	14	1	16	1	G=0 p=0.99 p=1	0	NA	NA	
sbd105	20	19	16	16	14	18	4	16	1	G=1.75 p=0.19 p=0.35	20	14	14	G=0.06 p=0.80
P115	20	12	10	16	12	16	6	16	1	G=3.48 p=0.06 p=0.11	20	20	2	G=11.32 p=0.01
RD31	20	16	9	15	3	14	2	11	2	G=1.73 p=0.19 p=0.35	20	17	11	G=0.78 p=0.38 p=0.48
DG2	24	22	15	23	12	20	12	15	3	G=1.87 p=0.17	20	15	12	G=0 p=1
ED5780	24	17	11	17	11	15	2	20	1	G=0.69 p=0.40 p=1	20	3	5	G=0.35 p=0.55 p=0.68
Cha7	20	19	14	18	14	20	1	18	0	G=1.41, p=0.23 p=1	20	7	0	G=8.29 p=0.01 p=0.03
Dl-BX12	20	20	16	15	13	16	0	18	0	G=0 p=1 p=1	20	2	3	G=0.39 p=0.53 p=0.65
H-B79	20	17	12	17	12	13	0	15	0	G=0 p=1 p=1	20	10	5	G=1.06 p=0.30
BSC43	24	18	12	19	11	21	1	19	0	G=1.51 p=0.22 p=1	20	5	2	G=1.16 p=0.28 p=0.39
e-R1	20	16	14	14	12	18	4	19	0	G=5.57 p=0.02 p=0.12	20	8	0	G=8.26 p=0.01 p=0.03
e-GC3	20	19	14	17	13	17	0	13	1	G=2.01 p=0.16 p=1	20	5	8	G=0.64 p=0.42
ED6096	20	16	13	20	17	15	0	18	8	G=10.12 p=0.01 p=0.02	0	NA	NA	

mel/bal		*mel*/*Df*		*sim*/bal		*sim*/*Df*					Sperm Assay			
Df(3R)	N	Crt	Cop	Crt	Cop	Crt	Cop	Crt	Cop	G test/Fishers	N	bal	*Df*	G test/Fisher's
BSC56	20	15	12	19	14	19	3	19	4	G=0.20 p=0.65 p=0.69	0	NA	NA	
BSC137	20	18	10	19	11	17	0	17	3	G=4.29 p=0.04 p=0.24	0	NA	NA	
BSC489	13	11	10	13	13	12	1	11	0	G=1.55, p=0.21 p=0.48	0	NA	NA	
Exel6196	20	14	9	15	9	17	2	16	0	G=2.91, p=0.09 p=0.49	0	NA	NA	
Exel6197	20	16	11	17	12	13	0	18	2	G=2.90, p=0.09 p=0.49	0	NA	NA	
ED6187	20	15	13	16	13	13	0	17	0	G=0 p=1 p=1	32	5	24	G=8.36 p=0.01
crb87-5	20	14	13	10	6	16	5	18	1	G=1.71 p=0.19 p=0.38	60	18	11	G=0.34 p=0.56
ED6220	20	16	8	19	13	16	4	19	9	G=0.22 p=0.64	0	NA	NA	
BSC461	20	13	8	14	9	17	4	19	5	G=0.01 p=0.93 p=1	0	NA	NA	
Exel6202	20	17	13	17	13	17	2	18	1	G=0.41 p=0.52 p=1	0	NA	NA	
Exel6203	20	9	6	13	10	13	0	13	1	G=1.85 p=0.17 p=1	20	3	4	G=0.06 p=0.81 p=1
BSC321	20	15	13	18	14	17	1	18	1	G=0.01 p=0.97 p=1	20	7	6	G=0.11 p=0.74
BSC140	20	16	14	16	13	17	0	20	2	G=2.83 p=0.09 p=0.49	0	NA	NA	
Espl3	20	16	10	18	11	15	4	16	1	G=2.12 p=0.15 p=0.36	20	11	6	G=1.68 p=0.20

Df(3R)	N	*mel*/bal		*mel*/*Df*		*sim*/bal		*sim*/*Df*			Sperm Assay			
		Crt	Cop	Crt	Cop	Crt	Cop	Crt	Cop	G test/Fishers	N	bal	*Df*	G test/Fisher's
BSC497	12	10	7	10	10	10	1	7	3	G=2.82 p=0.09 p=0.21	0	NA	NA	
IR16	20	15	9	16	12	15	5	16	9	G=0.23 p=0.63	0	NA	NA	
3450	20	14	9	16	10	16	0	20	2	G=2.68 p=0.10 p=0.49	0	NA	NA	
BSC547	20	15	2	20	1	12	0	16	0	G=0 p=1 p=1	0	NA	NA	
L127	44	31	26	39	25	28	3	29	1	G=0.89 p=0.35 p=0.39	20	3	3	G=0.01 p=0.97 p=1
BSC620	20	16	11	12	11	15	2	18	0	G=4.53 p=0.04 p=0.21	20	4	1	G=1.59 p=0.21 p=0.34
B81	20	18	13	19	11	17	11	20	3	G=3.03 p=0.08	20	16	4	G=3.34 p=0.07
Lines that tested previously identified female receptivity genes														
dissatisfaction														
(2L)cl-h3	28	21	19	24	19	20	6	22	4	G=0.34 p=0.56	20	10	13	G=0.21 p=0.65
PBac	20	19	17	18	17	18	1	16	0	G=1.57 p=0.21 p=1	20	3	2	G=0.18, p=0.68 p=1
spinster														
(2R)Jp1	20	16	15	18	14	13	5	15	2	G=1.32 p=0.25 p=0.16	20	5	5	G=0.01 p=0.93 p=1
P-insert	32	24	18	28	24	21	3	29	4	G=0.05 p=0.81 p=1	21	8	10	G=0.01 p=0.92

(Table 2.3), indicating that *D. melanogaster* males courted the hybrid females as often as the controls, and the deficiencies as often as the balancers.

For each line, I also compared the courtship latency (time until courtship began) by the *D. melanogaster* male towards the female of each of the four genotypes with use of a two-way ANOVA ($p < 0.05$). Six lines had a significant main effect of species (Table 2.4). For all the lines except for *Df*(2R)Jp1, *D. melanogaster* males courted the hybrid females significantly more rapidly than the pure-species *D. melanogaster* control females.

Two lines had a significant main effect of genotype (Table 2.4). For the line with a mutation in *spinster* P{ry[+t7.2]=neoFRT}43D, *D. melanogaster* males courted the females with a balancer chromosome significantly more rapidly than the females with the deficiency chromosome. For line *Df*(3R)ED5644 *D. melanogaster* males courted the females with a deficiency chromosome significantly quicker than the females with the balancer chromosome.

Four lines had a significant interaction effect: *Df*(3R)ED5780: $F(1,65)=10.51$, $p=0.002$; *Df*(3R)Exel6196: $F(1,59)=4.18$, $p=0.045$; *Df*(3R)ME15: $F(1,61)=6.63$, $p=0.012$; and *Df*(3R)crb87-5: $F(1,54)=4.64$, $p=0.036$. The time it took *D. melanogaster* males to start courting the females did not just depend on the chromosome that they inherited (balancer of deficiency) but also the background it was present in (pure species of hybrid).

Effect of balancer

Of the 51 lines that I used, there were 11 different balancers and 6 of the 11 were able to be assessed (Table 2.5). For both the 45-minute observational assay

Table 2.3: *D. melanogaster* male courtship occurrence toward the four types of females.

Lines that spanned 3R

Deficiency	*mel*/bal	*mel*/*Df*	*sim*/bal	*sim*/*Df*	G test
Df(3R)Me15	17	18	17	13	G=0.43, p=0.514
Df(3R)ED5156	20	20	21	19	G=0.05, p=0.823
Df(3R)Exel6144	17	17	20	15	G=0.35, p=0.552
Df(3R)ED5177	19	17	17	16	G=0.01, p=0.916
Df(3R)BSC464	18	18	14	19	G=0.40, p=0.528
Df(3R)Tpl10	4	6	5	7	G=0.01, p=0.937
Df(3R)ED5330	16	17	15	18	G=0.06, p=0.805
Df(3R)BSC38	17	18	16	16	G=0.01, p=0.907
Df(3R)M-Kx1	12	16	17	18	G=0.21, p=0.651
Df(3R)T-32	13	14	20	14	G=0.69, p=0.406
Df(3R)ry85	16	16	16	17	G=0.01, p=0.903
Df(3R)ED5644	25	26	27	26	G=0.04, p=0.844
Df(3R)ED5644	13	16	11	16	G=0.09, p=0.757
Df(3R)BSC471	17	15	14	16	G=0.26, p=0.611
Df(3R)sbd105	19	16	18	16	G=0.01, p=0.911
Df(3R)P115	12	16	16	16	G=0.31, p=0.58
Df(3R)RD31	16	15	14	11	G=0.11, p=0.743
Df(3R)DG2	22	23	20	15	G=0.54, p=0.463
Df(3R)ED5780	17	17	15	20	G=0.35, p=0.524
Df(3R)Cha7	19	18	20	18	G=0.01, p=0.912
Df(3R)Dl-BX12	20	15	16	18	G=0.70, p=0.401
Df(3R)H-B79	17	17	13	15	G=0.08, p=0.779
Df(3R)BSC43	18	19	21	19	G=0.11, p=0.736
Df(3R)e-R1	16	14	18	19	G=0.15, p=0.703
Df(3R)e-GC3	19	17	17	13	G=0.10, p=0.752
Df(3R)ED6096	16	20	15	18	G=0.01, p=0.933
Df(3R)BSC56	15	19	19	19	G=0.25, p=0.618
Df(3R)BSC137	18	19	17	17	G=0.01, p=0.909
Df(3R)BSC489	11	13	12	11	G=0.19, p=0.664
Df(3R)Exel6196	14	15	17	16	G=0.07, p=0.799
Df(3R)Exel6197	16	17	13	18	G=0.28, p=0.599
Df(3R)ED6187	15	16	13	17	G=0.18, p=0.692
Df(3R)crb87-5	14	10	16	18	G=0.72, p=0.397
Df(3R)ED6220	16	19	16	19	G=0, p=1
Df(3R)BSC461	13	14	17	19	G=0.01, p=0.942
Df(3R)Exel6202	17	17	17	18	G=0.01, p=0.906
Df(3R)Exel6203	9	13	13	13	G=0.40, p=0.528
Df(3R)BSC321	15	18	17	18	G=0.66, p=0.797
Df(3R)BSC140	16	16	17	20	G=0.11, p=0.737

Deficiency	*mel*/bal	*mel*/*Df*	*sim*/bal	*sim*/*Df*	G test
Df(3R)Espl3	16	18	15	16	G=0.01, p=0.915
Df(3R)BSC497	10	10	10	7	G=0.29, p=0.591
Df(3R)IR16	15	16	15	16	G=0, p=1
Df(3R)3450	14	16	16	20	G=0.03, p=0.857
Df(3R)BSC547	15	20	12	16	G=0, p=1
Df(3R)L127	31	39	28	29	G=0.30, p=0.587
Df(3R)BSC620	16	12	15	18	G=0.83, p=0.362
Df(3R)B81	18	19	17	20	G=0.05, p=0.816
Lines that tested previously identified female receptivity genes					
dissatisfaction					
Df(2L)cl-h3	21	24	20	22	G=0.01, p=0.929
PBac{w[+mC]=WH}	19	18	18	16	G=0.01, p=0.893
spinster					
Df(2R)Jp1	16	18	13	15	G=0.01, p=0.961
P{ry[+t7.2]=neoFRT}43D	24	28	21	29	G=0.36, p=0.546

Table 2.4: *D. melanogaster* courtship latency toward the four types of females.

Lines that spanned 3R

Df(3R)	*mel*/bal mean/SD	*mel*/*Df* mean/SD	*sim*/bal mean/SD	*sim*/*Df* mean/SD	Species	Genotype	Interaction
ME15	15.87/12.3	9.99/8.26	12.07/7.47	19.33/12.39	F(1,61)=1.18 p=.28	F(1,61)=0.07 p=.78	F(1,61)=6.63 p=.01
ED5156	13.17/9.87	14.77/8.11	10.08/8.03	6.56/4.02	F(1,76)=10.36 p=.01	F(1,76)=0.29 p=.59	F(1,76)=2.12 p=.15
Exel6144	13.48/12.1	12.52/11.2	15.02/10.9	8.22/5.56	F(1,65)=0.30 p=.59	F(1,65)=2.38 p=.13	F(1,65)=1.35 p=.25
ED5177	13.02/9.36	11.44/9.56	11.16/8.04	10.42/6.49	F(1,65)=0.49 p=.49	F(1,65)=0.32 p=.57	F(1,65)=0.04 p=.84
BSC464	11.12/11.0	8.43/5.70	10.17/6.18	6.95/5.79	F(1,65)=0.44 p=.51	F(1,65)=2.60 p=.11	F(1,65)=0.02 p=.88
Tpl10	9.83/7.33	8.37/3.70	13.71/14.9	17.81/12.41	F(1,18)=2.05 p=.17	F(1,18)=0.08 p=.78	F(1,18)= 0.36 p=.56
ED5330	13.57/8.95	14.16/10.9	14.87/12.8	10.95/7.98	F(1,62)=0.14 p=.71	F(1,62)=0.43 p=.51	F(1,62)=0.80 p=.38
BSC38	18.56/11.8	11.62/10.7	11.79/8.10	14.12/11.37	F(1,63)=0.67 p=.41	F(1,63)=0.79 p=.38	F(1,63)=3.18 p=.08
M-Kx1	14.54/8.94	12.86/9.63	9.64/7.96	9.48/7.43	F(1,59)=3.69 p=.06	F(1,59)=.183 p=.67	F(1,59)=0.12 p=.73
T-32	18.39/13.4	14.1/9.65	12.22/9.95	13.63/7.60	F(1,57)=1.55 p=.26	F(1,57)=0.29 p=.59	F(1,57)=1.15 p=.29
ry85	14.94/13.7	16.77/11.6	11.61/6.41	7.28/4.14	F(1,61)=7.11 p=.01	F(1,61)=0.27 p=.61	F(1,61)=1.64 p=.21
ED5644	20.53/15.3	14.72/11.7	12.54/12.9	8.74/5.00	F(1,100)=9.05 p=.01	F(1,100)=4.3 p=.04	F(1,100)=0.19 p=.66
ED5664	13.51/8.79	17.01/12.6	18.39/11.2	12.85/9.15	F(1,52)=0.20 p=.90	F(1,51)=0.13 p=.72	F(1,52)=2.50 p=.12
BSC471	19.62/5.38	23.17/13.7	10.49/6.82	10.50/9.85	F(1,58)=8.63 p=.01	F(1,58)=0.23 p=.63	F(1,58)=0.23 p=.63
sbd105	15.14/10.10	14.12/9.99	14.75/9.14	11.57/8.78	F(1,65)=0.41 p=.53	F(1,65)=0.83 p=.36	F(1,65)=0.22 p=.64
P115	14.06/9.90	11.57/7.72	16.59/12.5	11.54/7.41	F(1,56)=0.25 p=.62	F(1,56)=2.29 p=.13	F(1,59)=0.26 p=.61
RD31	11.26/4.90	11.67/11.1	14.08/10.90	15.11/10.58	F(1,52)=1.47 p=.23	F(1,52)=0.08 p=.78	F(1,52)=0.02 p=.90
DG2	12.18/8.28	11.90/7.50	12.61/10.9	12.85/9.97	F(1,76)=0.11 p=.74	F(1,76)=0 p=.99	F(1,76)=0.02 p=.90
ED5780	14.16/8.32	10.11/4.55	7.16/6.37	14.29/8.25	F(1,65)=0.67 p=.42	F(1,65)=0.79 p=.37	F(1,65)=10.50 p=.01
Cha7	13.75/11.00	12.04/11.2	14.26/11.50	9.76/9.98	F(1,71)=0.12 p=.73	F(1,71)=1.50 p=.22	F(1,71)=0.30 p=.58
Dl-BX12	14.15/12.40	13.04/8.38	14.54/9.21	15.2/12.77	F(1,65)=0.30 p=.58	F(1,65)=0 p=.99	F(1,65)=0.16 p=.69
H-B79	15.26/10.70	12.75/7.98	12.83/12.10	12.76/8.36	F(1,58)=0.23 p=.63	F(1,58)=0.27 p=.61	F(1,58)=0.24 p=.63
BSC43	13.71/11.30	15.47/11.00	13.12/5.67	12.28/8.53	F(1,73)=0.80 p=.37	F(1,73)=0.05 p=.83	F(1,73)=0.37 p=.54
e-R1	17.2/12.67	12.25/10.20	12.43/8.23	11.98/9.07	F(1,63)=1.04 p=.31	F(1,63)=1.20 p=.28	F(1,63)=0.84 p=.36

Df(3R)	*mel*/bal mean/SD	*mel*/*Df* mean/SD	*sim*/bal mean/SD	*sim*/*Df* mean/SD	Species	Genotype	Interaction
e-GC3	18.77/11.70	12.14/12.40	14.46/11.3	14.98/11.57	F(1,62)=0.06 p=.80	F(1,62)=1.10 p=.30	F(1,62)=1.49 p=.23
ED6096	13.73/10.50	13.37/10.50	18.80/13.8	14.09/10.72	F(1,65)=1.11 p=.30	F(1,65)=0.85 p=.36	F(1,65)=0.63 p=.43
BSC56	11.15/10.90	12.25/9.01	11.19/8.80	11.21/9.23	F(1,66)=0.05 p=.83	F(1,66)=0.06 p=.81	F(1,66)=0.06 p=.81
BSC137	15.04/12.40	14.21/9.61	20.76/14.10	11.57/8.61	F(1,67)=0.33 p=.57	F(1,67)=3.40 p=.07	F(1,67)=2.39 p=.13
BSC489	8.56/7.29	8.82/9.48	9.85/6.31	13.61/8.11	F(1,43)=1.72 p=.20	F(1,43)=0.75 p=.36	F(1,43)=0.57 p=.46
Exel6196	9.89/11.18	17.76/13.50	16.03/10.80	12.70/7.29	F(1,59)=0.04 p=.84	F(1,59)=0.67 p=.13	F(1,59)=4.18 p=.04
Exel6197	12.96/9.46	16.76/11.9	12.57/11.3	13.13/9.52	F(1,60)=0.57 p=.45	F(1,60)=0.67 p=.42	F(1,60)=0.37 p=.54
ED6187	18.96/15.00	10.00/6.85	13.23/6.68	11.19/12.75	F(1,57)=0.63 p=.43	F(1,57)=3.71 p=.06	F(1,57)=1.47 p=.23
crb87-5	14.82/10.00	7.69/6.40	11.11/10.10	14.76/9.36	F(1,54)=0.45 p=.51	F(1,54)=0.48 p=.49	F(1,54)=4.64 p=.04
ED6220	9.59/7.93	10.04/5.66	11.15/8.61	16.39/14.17	F(1,66)=2.87 p=.09	F(1,66)=1.48 p=.23	F(1,66)=1.05 p=.31
BSC461	8.72/5.32	12.92/9.52	10.89/7.45	13.96/10.03	F(1,59)=0.56 p=.46	F(1,59)=2.85 p=.10	F(1,59)=0.07 p=.79
Exel6202	10.17/8.77	9.81/6.41	12.73/8.78	12.87/10.16	F(1,65)=1.81 p=.18	F(1,65)=0.003 p=.96	F(1,65)=0.01 p=.91
Exel6203	12.90/6.07	15.21/12.00	17.78/12.00	13.07/8.93	F(1,46)=0.22 p=.64	F(1,46)=0.16 p=.69	F(1,46)=1.40 p=.24
BSC321	13.43/10.70	12.55/10.80	19.61/17.60	15.77/13.89	F(1,64)=2.02 p=.16	F(1,64)=0.51 p=.48	F(1,64)=0.20 p=.66
BSC140	17.21/9.17	15.71/8.96	14.74/8.15	9.55/5.73	F(1,65)=4.99 p=.03	F(1,65)=3.01 p=.09	F(1,65)=0.91 p=.34
Espl3	9.70/8.44	14.78/8.44	11.28/6.85	11.50/8.96	F(1,61)=0.12 p=.73	F(1,61)=1.22 p=.27	F(1,61)=1.02 p=.32
BSC497	10.38/7.41	9.88/7.41	15.67/9.12	12.21/11.44	F(1,33)=1.88 p=.18	F(1,33)=0.49 p=.49	F(1,33)=0.27 p=.61
IR16	12.99/11.90	12.14/12.00	13.8/10.20	12.56/6.89	F(1,58)=0.05 p=.82	F(1,58)=0.16 p=.70	F(1,58)=0.01 p=.94
3450	10.24/5.38	9.49/7.98	12.82/10.00	8.03/6.95	F(1,62)=0.08 p=.77	F(1,62)=2.05 p=.16	F(1,62)=2.05 p=.30
BSC547	12.20/9.24	18.21/11.6	16.84/11.4	15.14/11.76	F(1,59)=0.08 p=.78	F(1,59)=0.57 p=.45	F(1,59)=1.84 p=.18
L127	13.16/10.30	12.04/9.53	15.47/12.20	11.31/9.71	F(1,123)=0.18 p=.67	F(1,123)=2.01 p=.16	F(1,123)=0.67 p=.42
BSC620	11.03/6.12	11.78/11.70	8.68/5.98	9.47/9.57	F(1,57)=1.12 p=.29	F(1,57)=0.12 p=.73	F(1,57)=0.001 p=.99
B81	10.23/8.93	12.66/11.5	12.33/10.40	11.94/12.00	F(1,70)=0.07 p=.79	F(1,70)=0.16 p=.69	F(1,70)=0.31 p=.58

Lines that tested previously identified female receptivity genes

Df	*mel*/bal mean/SD	*mel*/*Df* mean/SD	*sim*/bal mean/SD	*sim*/*Df* mean/SD	Species	Genotype	Interaction
dissatisfaction							
(2L)cl-h3	12.46/10.1	12.47/9.34	14.26/7.14	9.54/7.84	F(1,59)=0.06 p=.80	F(1,59)=1.09 p=.30	F(1,59)=1.11 p=.30
PBac	7.91/4.90	11.61/6.61	10.87/11.7	9.09/6.96	F(1,67)=0.01 p=.91	F(1,67)=0.26 p=.61	F(1,67)=2.11 p=.15
spinster							
(2R)Jp1	9.19/10.34	10.01/7.74	19.38/12.6	14.72/7.84	F(1,58)=9.13 p=.004	F(1,58)=.61 p=.44	F(1,58)=1.23 p=.27
P-insert	9.84/5.49	13.32/9.83	10.21/5.25	12.59/6.77	F(1,98)=0.01 p=.90	F(1,98)=4.11 p=.04	F(1,98)=0.15 p=.70

Table 2.5: The effect of the balancer chromosome on female mating behaviour with *D. melanogaster* males.

Df(3R)	Balancer	Court	Cop	Prop mated	Balancer mean	Balancer SD	N	sperm	Prop	Balancer mean	Balancer SD
RD31	In(3R)C	14	2	0.14286	NA	NA	20	17	0.85	NA	NA
ME15	MKRS	17	1	0.058823529	0.0919	0.0468	20	13	0.65	0.65	0
ry85	MKRS	16	2	0.125			20	13	0.65		
T-32	MRS	20	4	0.2	NA	NA	0	NA	NA	NA	NA
P115	TM1	16	6	0.375	NA	NA	20	20	1	NA	NA
BSC38	TM2	16	1	0.0625	0.1430	0.2104	0	NA	NA	0.3610	0.2595
DG2	TM2	20	12	0.6			20	15	0.75		
ED5780	TM2	15	2	0.1333			20	3	0.15		
H-B79	TM2	13	0	0			20	10	0.5		
BSC43	TM2	21	1	0.04762			20	5	0.25		
BSC56	TM2	19	3	0.157894737			0	NA	NA		
ED6187	TM2	13	0	0			32	5	0.15625		
Tpl10	TM3	5	3	0.6	0.3759	0.1749	0	NA	NA	0.54	0.2074
M-Kx1	TM3	17	5	0.2941			32	16	0.5		
IR16	TM3	15	5	0.3333			0	NA	NA		
B81	TM3	17	11	0.6471			20	16	0.8		
e-R1	TM3	18	4	0.2222			20	8	0.4		
sbd105	TM3	18	4	0.2222			20	14	0.7		
crb87-5	TM3	16	5	0.3125			60	18	0.3		
L127	TM6	28	3	0.1071	NA	NA	20	3	0.15	NA	NA
Exel6144	TM6b	20	0	0	0.0259	0.0477	0	NA	NA	0.2125	0.1109
Cha7	TM6b	20	1	0.05			20	7	0.35		
Dl-BX12	TM6b	16	0	0			20	2	0.1		
E-gc3	TM6b	17	0	0			20	5	0.25		
BSC137	TM6b	17	0	0			0	NA	NA		
Exel6196	TM6b	17	2	0.1176			0	NA	NA		

Df(3R)	Balancer	Court	Cop	Prop mated	Balancer mean	Balancer SD	N	sperm	Prop	Balancer mean	Balancer SD
Exel6197	TM6b	13	0	0			0	NA	NA		
Exel6202	TM6b	17	2	0.1176			0	NA	NA		
Exel6203	TM6b	13	0	0			20	3	0.15		
BSC140	TM6b	17	0	0			0	NA	NA		
3450	TM6b	16	0	0			0	NA	NA		
ED5156	TM6c	21	2	0.0952	0.0915	0.0906	20	7	0.35	0.35	0.1433
ED5177	TM6c	17	1	0.0588			0	NA	NA		
BSC464	TM6c	14	0	0			0	NA	NA		
ED5330	TM6c	15	0	0			20	7	0.35		
ED5644	TM6c	27	3	0.1111			36	18	0.5		
ED5664	TM6c	11	0	0			20	3	0.15		
BSC471	TM6c	14	1	0.0714			0	NA	NA		
ED6098	TM6c	15	0	0			0	NA	NA		
BSC489	TM6c	12	1	0.0833			0	NA	NA		
ED6220	TM6c	16	4	0.25			0	NA	NA		
BSC461	TM6c	17	4	0.2353			0	NA	NA		
BSC321	TM6c	17	1	0.0588			20	7	0.35		
BSC497	TM6c	10	1	0.1			0	NA	NA		
BSC547	TM6c	12	0	0			0	NA	NA		
BCS620	TM6c	15	2	0.1333			20	4	0.2		
Espl3	TM6c	15	4	0.2667			20	11	0.55		
(2L)cl-h3	SM6b	20	6	0.3	NA	NA	20	10	0.5	NA	NA
PBac	CyO	18	1	0.0556	0.1943	0.1705	20	3	0.15	0.2603	0.1158
(2R)Jp1	CyO	13	5	0.3846			20	5	0.25		
P-insert	CyO	21	3	0.1429			21	8	0.3810		

and the 24 hour sperm assay, a significant difference was found among the balancers for copulation occurrence ($F(5,40)=7.11$, $p=0.0001$; $F(5,20)=2.79$, $p=0.05$, respectively). A Tukey's post hoc test showed that the significance was driven by the TM3 balancer females mating more often than females with the TM2, TM6b, and TM6c balancers, and not by females with a specific balancer mating less often compared to the others. When TM3 data was removed, no significant difference was found for the behaviour or sperm assay results ($F(4,34)=1.78$, $p=0.16$; $F(4,16)=2.44$, $p=0.09$, respectively).

To investigate whether deficiency chromosomes that had been maintained over a particular balancer caused females to be more or less likely to copulate, I made the same comparisons as for the *sim*/bal (above) by grouping the *sim/Df* female scores according to which balancer chromosome was used in that stock. No significant difference was found for both the 45-minute observational assay ($F(5,40)=0.94$, $p=0.45$) and for the 24-hour sperm assay ($F(5,20)=0.49$,$p=0.79$). However, the *sim/Df* females from the TM3 lines had the greatest mean copulation occurrence (Table 2.6).

Discussion

I mapped the right arm of the third chromosome for gene(s) that cause female *D. simulans* to reject courting male *D. melanogaster*. With use of deficiency mapping, I located five regions that influence this behaviour. All five of these regions reside in areas of low recombination: within the interspecific inversion polymorphism, near the centromere, and near the telomere.

Table 2.6: The effect of the genetic background of the deficiency chromosome on female mating behaviour with *D. melanogaster* males.

Df(3R)	Balancer	Court	Cop	Prop mated	Balancer mean	Balancer SD	N	sperm	Prop	Balancer mean	Balancer SD
RD31	In(3R)C	11	20.18	1818182	NA	NA	20	11	0.55	NA	NA
ME15	MKRS	13	0	0	0.1176	0.1664	20	3	0.15	0.4	0.3536
ry85	MKRS	17	4	0.2353			20	13	0.65		
T-32	MRS	14	1	0.0714	NA	NA	0	NA	NA	NA	NA
P115	TM1	16	1	0.0625	NA	NA	20	2	0.1	NA	NA
BSC38	TM2	16	3	0.1875	0.0926	0.1016	0	NA	NA	0.39	0.2725
DG2	TM2	15	3	0.2			20	12	0.6		
ED5780	TM2	20	1	0.05			20	5	0.25		
H-B79	TM2	15	0	0			20	5	0.25		
BSC43	TM2	19	0	0			20	2	0.1		
BSC56	TM2	19	4	0.2105			0	NA	NA		
ED6187	TM2	17	0	0			32	24	0.75		
Tpl10	TM3	7	6	0.8571	0.2411	0.3351	0	NA	NA	0.2854	0.2620
M-Kx1	TM3	18	0	0			32	11	0.3438		
IR16	TM3	16	9	0.5625			0	NA	NA		
B81	TM3	20	3	0.15			20	4	0.2		
e-R1	TM3	19	0	0			20	0	0		
sbd105	TM3	16	1	0.0625			20	14	0.7		
crb87-5	TM3	18	1	0.0556			60	11	0.1833		
L127	TM6	29	1	0.0345	NA	NA	20	3	0.15	NA	NA
Exel6144	TM6b	15	0	0	0.0634	0.0585	0	NA	NA	0.1875	0.1652
Cha7	TM6b	18	0	0			20	0	0		
Dl-BX12	TM6b	18	0	0			20	3	0.15		
E-gc3	TM6b	13	1	0.0769			20	8	0.4		
BSC137	TM6b	17	3	0.1765			0	NA	NA		

Df(3R)	Balancer	Balancer Court	Cop	Prop mated	Balancer mean	Balancer SD	N	sperm	Prop	Balancer mean	SD
Exel6196	TM6b	16	0	0			0	NA	NA		
Exel6197	TM6b	18	2	0.1111			0	NA	NA		
Exel6202	TM6b	18	1	0.0556			0	NA	NA		
Exel6203	TM6b	13	1	0.0769			20	4	0.2		
BSC140	TM6b	20	2	0.1			0	NA	NA		
3450	TM6b	20	2	0.1			0	NA	NA		
ED5156	TM6c	19	1	0.0526	0.1293	0.1717	20	6	0.3	0.3389	0.1550
ED5177	TM6c	16	1	0.0625			0	NA	NA		
BSC464	TM6c	19	0	0			0	NA	NA		
ED5330	TM6c	18	0	0			20	9	0.45		
ED5644	TM6c	26	1	0.0385			36	17	0.4722		
ED5664	TM6c	16	2	0.125			20	10	0.5		
BSC471	TM6c	16	1	0.0625			0	NA	NA		
ED6098	TM6c	18	8	0.4444			0	NA	NA		
BSC489	TM6c	11	0	0			0	NA	NA		
ED6220	TM6c	19	9	0.4737			0	NA	NA		
BSC461	TM6c	19	5	0.2632			0	NA	NA		
BSC321	TM6c	18	1	0.0556			20	6	0.3		
BSC497	TM6c	7	3	0.4286			0	NA	NA		
BSC547	TM6c	16	0	0			0	NA	NA		
BCS620	TM6c	18	0	0			20	1	0.05		
Espl3	TM6c	16	1	0.0625			20	6	0.3		
(2L)cl-h3	SM6b	22	4	0.1818	NA	NA	20	13	0.65	NA	NA
PBac	CyO	16	0	0	0.0904	0.0783	20	2	0.1	0.2754	0.1894
(2R)Jp1	CyO	15	2	0.1333			20	5	0.25		
P-insert	CyO	29	4	0.1379			21	10	0.4762		

Male courtship behaviour

By comparing courtship occurrence of the four genotypes (*mel*/bal, *mel*/*Df*, *sim*/bal, and *sim*/*Df*) I determined that all four types of females were equally courted by *D. melanogaster* males. Although Coyne (1996) found that at least one locus on the right arm of the third chromosome contributes to differences in *D. simulans* and *D. melanogaster* female cuticular hydrocarbons (CHCs), which are the *Drosophila* contact pheromones, the difference in the CHC profile may not be large enough to be detected by *D. melanogaster* males, or may fall within one of the regions I was unable to map. It is also possible that the effect of the right arm is due to many genes of small effect; each one individually may not be sufficient to alter the pheromone profile in a way that reduces a male's attractiveness to that female. Alternatively, since *D. melanogaster* males will court *D. simulans* females, it is possible that *D. melanogaster* males do not use CHCs as a cue for determining appropriate mating partners (Jamart *et al.* 1993). However, I did find that *D. melanogaster* males courted the hybrid females significantly more rapidly than the *D. melanogaster* females in five of the 51 lines. This may be due to a lack of inhibitory signals in the hybrids. Billeter *et al.* (2009) found that conspecific females without any CHCs were courted more by *D. melanogaster* males than wild-type females, and they suggested that this may be due to a female adapted CHC profile used to slow male courtship in order for females to better assess their potential mates.

For line *Df*(2R)Jp1, *D. melanogaster* males courted pure species *D. melanogaster* controls significantly faster than hybrids. This could potentially bias the female behaviour results. Although these females were courted at a slightly later

time, they were not courted less frequently, nor did they copulate less frequently within the 45 minute behaviour assay and no significant difference was found between the mating behaviour of the females.

For two of the lines, the females with different chromosomes (balancer or deficiency) were courted at different speeds by the *D. melanogaster* males. Only the line with the P-element insertion disrupting the *spinster* gene had a significant effect that could indicate that the timing of male courtship may bias the interpretation: it took significantly more time for *D. melanogaster* males to court females with the deficiency chromosome than females with the balancer chromosome. This could potentially bias the interpretation of the female receptivity results as *sim/Df* females would have less time to be receptive, however, no significant difference was found in the female mating behaviour.

Four lines had a significant interaction effect genotype and species (*Df*(3R)ME15, *Df*(3R)ED5780, *Df*(3R)Exel6196, and *Df*(3R)crb87-5). One line had a significant effect driven by differences in the rate at which the balancer genotypes were courted (*Df*(3R)ED5780), but in the opposite direction if it were to bias the interpretation of the female receptivity results (*mel*/bal were courted more slowly than *sim*/bal). One line had a significant effect due to the rate at which the deficiencies were courted (*Df*(3R)Exel6196), but in the opposite direction if it were to bias the interpretation of the female receptivity results (*mel/Df* were courted more slowly than *sim/Df).* Only two lines have a significant effect that could indicate that the timing of male courtship may bias the interpretation of the female receptivity results (*Df*(3R)ME15 and *Df*(3R)crb87-5), and only one of these lines

(*Df*(3R)ME15) was found to have significantly less copulation for the *sim/Df* genotype. Although these females were courted at a slightly later time, they were not courted less frequently, nor did they copulate less frequently within the 45 minute behaviour assay. It is therefore unlikely that the approximately 7 minute lag in male courtship of *sim/Df* compared to *sim*/bal females is responsible for the significant reduction in the 24 hour rate of copulation.

Effect of the balancer

A balancer effect was found on mating behaviour but not in the way that was predicted. If the breakpoints of a specific balancer disrupted a gene for interspecific female preference, it would act as a deficiency, uncover the *D. simulans* genome, and thus a reduction in mating of *sim*/bal females would be observed. This did not occur. However, I did find that hybrid females with TM3 balancer showed overall higher mating occurrence with *D. melanogaster* males compared to hybrid females with another balancer. This phenomenon is most likely not due to the balancer breakpoints, but instead be due to the genetic background in which the balancer chromosome was created because *sim/Df* females from the TM3 lines had a non-significant trend of increased mating as well. There is a strong genetic basis for mating propensity, and it has previously been shown that the propensity to mate with conspecific (Moehring and Mackay 2004) and heterospecific (Izquierdo *et al.* 1992) individuals can be readily altered in selection experiments. It is possible that the TM3 balancer originated out of a *D. melanogaster* line which contained alleles that, when in a hybrid background, interact with *D. simulans* alleles and cause that female to be more receptive to *D. melanogaster* males.

Mating

I tested two genes that had been previously identified as affecting female preference within *D. melanogaster*: *spinster* and *dissatisfaction*. When mutated to loss-of-function, these two genes greatly reduce a female's willingness to mate (Suzuki *et al.* 1997; Finley *et al.* 1997 respectively). When the only functioning copy of this gene in interspecies hybrids came from *D. simulans*, however, there was not a difference in those female's willingness to mate with *D. melanogaster* males. This may be because their effect is only present when the genes are disrupted and not when there is simply allelic variants, that the same genes have completely different functions in *D. simulans* and *D. melanogaster*, or that genes responsible for intraspecific female receptivity are not the same as those responsible for interspecific female receptivity. Since it was previously shown that genes influencing variation in mating behaviour are not the same ones that are identified through mutagenesis studies (Moehring and Mackay 2004), it seems that this explanation is the most likely, but also prevents a true assessment of whether *variation* in genes for intraspecific mating also influence interspecific mating.

I also tested most of the genomic regions that span the right arm of the third chromosome. For the 45 minute behavioural assay, no lines had a significant behavioural effect, however, four lines were considered 'potentially significant': lines *Df*(3R)P115, *Df*(3R)e-R1, *Df*(3R)B81, and *Df*(3R)M-Kx1. For the 24 hour sperm assay, 5 lines were considered significant: *Df*(3R)P115, *Df*(3R)e-R1, *Df*(3R)B81, *Df*(3R)ME15, and *Df*(3R)Cha7. Three lines were both potentially significant during the 45 minute behavioural assay and significant for the 24 hour sperm assay

(*Df*(3R)P115, *Df*(3R)e-R1, and *Df*(3R)B81). Therefore, these regions were considered to house genes for interspecific female preference. One line (*Df*(3R)M-Kx1) was considered potentially significant from the 45 minute assay and was not considered significant or potentially significant for the 24 hour sperm assay. In these regions, the genetic effect, if any, is likely mild – potentially detectable in early measures of courtship and copulation, but not over longer periods of time. Therefore, this region was not considered to house genes for interspecific female preference. Finally, two lines that were not significant based on measures during the 45 minute behaviour assay (*Df*(3R)ME15 and *Df*(3R)Cha7) were found to be significant during the 24 hour sperm assay. Both of these lines had extremely low overall mating of the interspecies hybrids (both *sim/Df* and *sim*/bal) within the 45 minute assay, and thus the effect of the deficiency was likely only able to be detected with a longer assay period that increased the overall number of matings (primarily for the *sim*/bal genotype in these cases). Therefore, these regions were considered to house genes for interspecific female preference.

Not all lines were able to be tested with the 24 hour sperm assay. I therefore compared the lines that were tested with this assay to their behaviour assay results to see whether it is likely that lines that were not significant for the behaviour assay may have been found to be significant if they had also been tested with the sperm assay. Only two lines that were tested with the 24 hour sperm assay had non-significant results for the behaviour assay (*Df*(3R)ME15 and *Df*(3R)Cha7, discussed above). Both of these lines had extremely low mating within the 45 minute assay time (both had one mating out of 20 assays for *sim*/bal and zero matings for *sim/Df,*

greatly limiting the power to detect significant differences within that assay period. Only two lines that did not have a 24 hour sperm assay test had a behaviour assay p-value under p=0.50 (*Df*(3R)BSC489: G=1.55, p=0.213; *Df*(3R)Exel6196: G=2.91 p=0.088) and in both cases a Fisher's Exact test was used (p=0.48 and p=0.49, respectively). Therefore, I would consider the small region uncovered by these overlapping deficiencies (94F3-95D8) as possibly requiring additional testing with sperm assays to confirm that it does not contribute to behavioural isolation.

By methodically testing the right arm of the third chromosome using deficiency mapping, I found five regions that contribute to the rejection of *D. melanogaster* males by *D. simulans* females (Figure 2.2). The third chromosome was chosen for these assays because it was found to be important in the genetic basis of behavioural isolation in multiple species pairs of *Drosophila* (Coyne 1992; Uenoyama and Inoue 1995; Carracedo *et al.* 1995; Doi *et al.* 2001; Noor *et al.* 2001, Moehring *et al.* 2004; Moehring *et al.* 2006; Sawamura *et al.* 2008), with a strong effect of the right arm in particular (Doi *et al.* 2001; Noor *et al.* 2001; Moehring *et al.* 2006; Sawamura *et al.* 2008). Furthermore, this arm houses the only interspecific inversion polymorphism between this species pair, and such inversions have been shown to be important in species isolation in *Drosophila* (Noor *et al.* 2001; Nanda and Singh 2001; Sawamura *et al.* 2008; Sousa-Nerves and Rosas 2010), other animals (Ayala *et al.* 2011), and plants (Lowry and Willis 2010).

Our identification of regions containing behavioural isolation genes provides yet another example of genes for species isolation residing in regions of low recombination. All five regions were found in areas of low recombination: one close

to the centromere (81F3-82F7), one close to the telomere (99F2-3Rt), and three within the interspecific inversion polymorphism (89B10-89E2; 91B2-91F1; 93B7-93C5). The presence of inversions in interbreeding populations is widespread and found in both plants (Lowry and Willis 2010) and animals (Nanda and Singh 2001; Noor *et al.* 2001). Within these inversions, the reduction of recombination creates co-adapted gene complexes (genes inherited together), which, over time, can maintain new mutational variants through genetic hitchhiking. This is supported with the finding that regions of higher recombination contain lower interspecific nucleotide divergence (Stevison *et al.* 2011). Genetic variants within the complexes can contain variants for local adaptation (Feder *et al.* 2011), influence assortative mating within species (Nanda and Singh 2001), and influence behavioural isolation between species, either directly or as a byproduct of natural selection (Lowry and Willis 2010; Noor *et al.* 2001).

Conclusions

Although this study did not identify individual genes contributing to behavioural isolation, further studies can readily be done using the same deficiency mapping methodology to fine-map these regions down to the gene level due to the multitude of genetic tools available in *D. melanogaster*. Additional deficiency lines can first be tested to reduce the number of candidate genes, then individual gene mutants can be tested using the same approach as for the deficiency mapping: the only functional copy in an interspecies hybrid will be that from *D. simulans*. The eventual identification of individual genes that influence female preference between these two species of *Drosophila* will allow a multitude of additional questions to be

addressed: What type of genes (morphological, neural, etc.) influence female mate choice? What selective pressure has shaped the evolution of these genes? Do the genes that isolate species from one another also affect mate preference within a species? Future identification of multiple genes for behavioural isolation in a variety of species pairs will allow us to understand if general trends underlie the genetic basis of behavioural isolation, speciation, and the maintenance of biodiversity.

References

Ayala, D., M. C. Fontaine, A. Cohuet, D. Fontenille, R. Vitalis, and F. Simard. 2011. Chromosomal inversions, natural selection and adaptation in the malaria vector

Balakireva, M., R. F. Stocker, N. Gendre, and J. F. Ferveur. 1998. *Voila*, a new *Drosophila* courtship variant that affects the nervious system: behavioral, neural, and genetic characterizations. J. Neurosci. 18:4335-4343.

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.

Bour, B. A., M. Chakravarti, J. M. West, and S. M. Abmayr. 2000. *Drosophila* SNS, a member of the immunoglobulin superfamily that is essential for myoblast fusion. Genes Dev. 14:1498-1511.

Carney, G. E. 2007. A rapid genome-wide response to *Drosophila melanogaster* social interactions. BMC Genomics 8:288.

Carracedo, M. C., R. Piñeiro, and P. Casares. 1995. Chromosomal substitution analysis of receptivity and sexual isolation in *Drosophila melanogaster* females. Heredity 75:541-546.

Coyne, J. A. 1989. Genetics of sexual isolation between two sibling species, *Drosophila simulans* and *Drosophila mauritiana*. Proc. Natl. Acad. Sci USA 86:5464-5468.

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

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

Coyne, J. A., and H. A. Orr. 1997. "Patterns of speciation in *Drosophila*" revisited. Evolution 51:295-303.

Davis, A. W., J. Roote, T. Morley, K. Sawamura, S. Herrmann, and M. Ashburner. 1996. Rescue of hybrid sterility in crosses between *D. melanogaster* and *D. simulans*. Nature 380:157-159.

Doi, M., M. Matsuda, M. Tomaru, H. Matsubayashi, and Y. Oguma. 2001. A locus for female discrimination behavior causing sexual isolation in *Drosophila*. Proc. Natl. Acad. Sci. USA 98:6714-6719.

Drosophila 12 Genomes Consortium. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. Nature 450:203-218.

Feder, J. L., R. Gejji, T. H. Powell, and P. Nosil. 2011. Adaptive chromosomal divergence driven by mixed geographic mode of evolution. Evolution 65:2157-2170.

Ferveur, J. F. 2005. Cuticular hydrocarbons: their evolution and roles in *Drosophila* pheromonal communication. Behav. Genet. 35:279-295.

Finley, K. D., B. J. Taylor, M. Milstein, and M. McKeown. 1997. *dissatisfaction*, a gene involved in sex-specific behavior and neural development of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 94:913-918.

Greenspan, R. J. 1995. Understanding the genetic construction of behavior. Sci. Am. 272:72-78.

Greenspan, R. J., and J. F. Ferveur. 2000. Courtship in *Drosophila*. Annu. Rev. Genet. 34:205-232.

Hall, J. C. 1994. The mating of a fly. Science 264:1702-1714.

Izquierdo, J. I., M. C. Carracedo, R. Piñeiro, and P. Casares. 1992. Response to selection for increased hybridization between *Drosophila melanogaster* females and *D. simulans* males. J. Hered. 83:100-104.

Jamart, J. A., M. C. Carracedo, and P. Casares. 1993. Sexual isolation between *Drosophila melanogaster* females and *D. simulans* males. Male mating propensities versus success in hybridization. Experientia 49:596-598.

Juni, N., and D. Yamamoto. 2009. Genetic analysis of *chaste*, a new mutation of *Drosophila melanogaster* characterized by extremely low female sexual receptivity. J. Neurogenet. 23:329-340.

Klahre, U., A. Gurba, K. Hermann, M. Saxenhofer, E. Bossolini, P. M. Guerin, and C. Kuhlemeier. 2011. Pollinator choice in *Petunia* depends on two major genetic Loci for floral scent production. Curr. Biol. 21:730-739.

Kronforst, M. R., L. G. Young, D. D. Kapan, C. McNeely, R. J. O'Neill, and L. E. Gilbert. 2006. Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. Proc. Natl. Acad. Sci. USA 103:6575-6580.

Kulkarni, S. J., A. F. Steinlauf, and J. C. Hall. 1988. The dissonance mutant of courtship song in *Drosophila melanogaster*: isolation, behavior and cytogenetics. Genetics 118:267-285.

Kyriacou, C. P., and J. C. Hall. 1980. Circadian rhythm mutations in *Drosophila melanogaster* affect short-term fluctutaions in the male's courtship song. Proc. Natl. Acad. Sci. USA 77:6729-6733.

Lowry, D. B., and J. H. Willis. 2010. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. PLoS ONE 8:e1000500.

Mackay, T. F., S. L. Heinsohn, R. F. Lyman, A. J. Moehring, T. J. Morgan, and S. M. Rollmann. 2005. Genetics and genomics of *Drosophila* mating behavior. Prac. Natl. Acad. Sci. USA 102:6622-6629.

Moehring, A. J., A. Llopart, S. Elwyn, J. A. Coyne, and T. F. Mackay. 2006. The genetic basis of prezygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to mating preference. Genetics 173:215-223.

Moehring, A. J., and T. F. Mackay. 2004. The quantitative genetic basis of male mating behavior in *Drosophila melanogaster*. Genetics 167:1249-1263.

Moehring, A. J., J. Li, M. D. Schug, S. G. Smith, M. deAngelis, T. F. Mackay, and J. A. Coyne. 2004. Quantitative trait loci for sexual isolation between *Drosophila simulans* and *D. mauritiana*. Genetics 167:1265-1274.

Moulin, B., T. Aubin, and J. M. Jallon. 2004. Why there is a one-way crossability between *D. melanogaster* and *D. simulans*? An ontogenic explanation. Genetica 120:285-292.

Nanda, P., and B. N. Singh. 2011. Effect of chromosome arrangements on mate recognition system leading to behavioral isolation in *Drosophila ananassae*. Genetica 139:273-279.

Narraway, C., J. Hunt, N. Wedell, and D. J. Hosken. 2010. Genotype-by-environment interactions for female preference. J. Evol. Biol. 23:2550-2557.

Noor, M. A., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. Prac. Natl. Acad. Sci. USA 98:12084-12088.

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

Peixoto, A. A., and J. C. Hall. 1998. Analysis of termperature-sensitive mutants reveals new genes involved in the courtship of *Drosophila*. Genetics 148:827-838.

Perotti, M. E., F. Cattaneo, M. E. Pasini, F. Vernì, and J. H. Hackstein. 2001. Male sterile mutant *casanova* gives clues to mechanisms of sperm-egg interactions in *Drosophila melanogaster*. Mol. Reprod. Dev. 60:248-259.

Presgraves, D. C., L. Balagopalan, S. M. Abmayr, and H. A. Orr. 2004. Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. Nature 423:715-719.

Ranz, J. M., D. Maurin, Y. S. Chan, M. von Grotthuss, L. W. Hillier, J. Roote, M. Ashburner, and C. M. Bergman. 2007. Principles of genome evolution in the *Drosophila melanogaster* species group. PLoS ONE 5:e152.

Ruedi, E. A., and K. A. Hughes. 2008. Natural genetic variation in complex mating behaviors of male *Drosophila melanogaster*. Behav. Genet. 38:424-436.

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

Sawamura, K., H. Zhi, K. Setoguchi, H. Yamada, T. Miyo, M. Matsuda, and Y. Oguma. 2008. Genetic analysis of female mating recognition between *Drosophila ananassae* and *Drosophila pallidosa*: application of interspecific mosaic genome lines. Genetica 133:179-185.

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

Stevison, L. S., K. B. Hoehn, and M. A. Noor. 2011. Effects of inversions on within- and between-species recombination and divergence. Genome Biol. Evol. 3:830-841.

Sturtevant, A. H. 1920. Genetic studies on *Drosophila simulans*. I. Introduction. Hybrids with *Drosophila melanogaster*. Genetics 5:488-500.

Suzuki, K., N. Juni, and D. Yamamoto. 1997. Enhanced mate refusal in female *Drosophila* induced by a mutation in the *spinster* locus. Appl. Entomol. Zool. 32:235-243.

Tamura, K., S. Subramanian, and S. Kumar. 2004. Temporal Patterns of Fruit Fly (*Drosophila*) evolution revealed by mutation clocks. Mol. Biol. Evol. 21:36-44.

Ting, C. T, A. Takahashi, and C. I. Wu. 2001. Incipient speciation by sexual isolation in *Drosophila*: concurrent evolution at multiple loci. Proc. Natl. Acad. USA 98:6709-6713.

Tweedie, S., M. Ashburner, K. Falls, P. Leyland, P. McQuilton, S. Marygold, G. Millburn, D. Osumi-Sutherland, A. Schroeder, R. Seal, H. Zhang, and The FlyBase Consortium. 2009. FlyBase: enhancing Drosophila Gene Ontology annotations. Nucleic Acids Research 37:D555-D559.

Uenoyama, T., and Y. Inoue. 1995. Genetic studies on premating isolation in *Drosophila simulans*. I. A *D. simulans* line highly crossable with *D. melanogaster*. Jpn. J. Genet. 70:365-371.

Chapter 3: Finer mapping of four regions previously identified to contain genes for interspecific female preference in *D. simulans*.

Introduction

Species are separated by barriers that reduce the gene flow between two or more interbreeding populations. These barriers have been divided into two functional categories: prezygotic and postzygotic. The genetic basis of postzygotic barriers, which act after fertilization, has been extensively investigated and led to the identification of specific genes that cause hybrid sterility or inviability (Masly *et al.* 2006; Lee *et al.* 2008; Bayes and Malik 2009; Mihola *et al.* 2009). Although much research has also investigated the genetic basis of prezygotic barriers (Wheeler 1991; Moehring *et al.* 2004; Gleason *et al.* 2005), which act prior to fertilization, progress in this field has advanced at a slower pace. Although both mechanisms function to maintain genetic boundaries between species, behavioural isolation is thought to be the first to evolve during the process of species divergence (Coyne and Orr 1997). Thus, the identification of the genetic variants that produce behavioural isolation will further our knowledge and understanding of the process of speciation.

The lack of advances in this area of research is due to a disconnect between the research question and the methods available to answer it. The most common method in genetics for locating genes that contribute to a quantitative trait such as behaviour is namely recombination mapping. This type of mapping requires crossing individuals from two reproductively isolated groups and producing fertile offspring. However, by definition, hybrids of two species are usually not fertile or

viable. Second, identifying the genes associated with a behaviour requires the location of multiple genes with different effect sizes, necessitating a repeatable measure of the behaviour, large sample sizes, and the availability of powerful genetic tools. Both of these primary obstacles can be overcome by the use of the *Drosophila* model. First, many *Drosophila* sister species are only partially isolated in a lab setting, producing viable and fertile hybrids (Coyne 1992). Second, *Drosophila melanogaster* has a large number of genetic tools available, including a sequenced genome and single gene mutants, which allows for easy identification of genes.

The mating behaviour of *D. melanogaster* follows a specific sequence of auditory signals, chemical signals, stereotypical behaviours, and cooperation of both sexes (Hall 1994). Male courtship behaviour begins with locating and orientating to a focal female. Once the male has identified his potential mate, he sends and receives chemical signals by tapping the female, performing unilateral wing vibrations, and contacting her genitalia with his proboscis. The wing vibration also sends auditory signals as the vibrations produce a species-specific courtship song (Hall 1994). The female has stereotypical behaviours of her own that communicate that she is receptive, which include reducing her locomotor behaviour, positioning herself correctly to the male, and presenting her abdomen by spreading her wings (Griffith and Ejima 2009). Finally, the male will attempt copulation by thrusting his genitalia towards hers (Hall 1994) and the female will allow for insertion by opening her genital plates (Greenspan 1995).

Mutagenesis studies have identified multiple genes that are important in the normal mating behaviour of both male and female *D. melanogaster*. These studies

have provided information on what types of genes can influence the construction of mating behaviour such as transcription factors (*fruitless*, Hall 1978; *dissatisfaction*, Finley *et al.* 1998) and where the genes for mating behaviour are expressed such as in the central nervous system (*spinster*, Suzuki *et al.* 1997; *dissatisfaction*, Finley *et al.* 1998; *chaste,* Juni and Yamamoto 2009; *voila*, Balakireva *et al.* 1998; *fruitless*, Hall 1978; *5-HT7* Becnel *et al.* 2011). However, these genes may not contribute to behavioural isolation for multiple reasons.

First, the vast majority of the genes identified in these studies were genes for male traits. However, in the case of species isolation, males have more often been found to court both con- and heterospecific females (Jamart *et al.* 1993). Females, on the other hand, will usually discriminate against heterospecific males (Noor *et al.* 2001; Moehring *et al.* 2004; Doi *et al.* 2001; Noor *et al.* 2001; Moehring *et al.* 2006; Sawamura *et al.* 2008; Carracedo *et al.* 1998; Uenoyama and Inoue 1995), and can easily prohibit copulation with unsuitable partners by presenting her ovipositor and making male genital insertion impossible, or simply by flying away (Hall 1994). Therefore, to identify a gene for behavioural isolation, attention must be paid to the genetic basis of female mating behaviours such as rejection of heterospecific males.

Second, mutation studies have been aimed at successfully identifying the genetic basis of conspecific mating behaviour. Although females of different species may focus on the same male traits (Immonen and Ritchie 2011), intra- and interspecific mating behaviour may not have the same genetic basis (Moehring and Mackay 2004; Gleason *et al.* 2005; Mackay *et al.* 2005; Immonen and Ritchie 2011).

Finally, mutation studies completely abolish the normal function of the gene in order to assess its affect on behaviour. This demonstrates that the gene is important for formation of the behaviour but it does not determine if the gene contributes to variation in the behaviour or if it differs between species. Therefore, to determine the genetic basis of behavioural isolation, studies must focus on female mating behaviour, concentrate on heterospecific interactions, and highlight natural variants that exist between species.

To date, no individual genes have been identified that influence interspecific female preference, although the trait has a clear heritable basis (Hall 1994). However, through the identification of genomic regions that contribute to the trait, two patterns have become apparent. First, many studies identify regions along the third chromosome in *D. melanogaster* (or the analogous chromosome in closely related species) as containing loci for behavioural isolation (Coyne 1992; Uenoyama and Inoue 1995; Carracedo *et al.* 1995; Doi *et al.* 2001; Noor *et al.* 2001, Moehring *et al.* 2004, 2006; Sawamura *et al.* 2008). When fine mapping was available, most regions were specifically located on the right arm of the third chromosome (Noor *et al.* 2001; Moehring *et al.* 2006; Sawamura *et al.* 2008). Second, loci for behavioural isolation are repeatedly found in areas of low recombination, which include regions near centromeres, telomeres, and within inversion polymorphisms (Noor *et al.* 2001; Nanda and Singh 2001; Sawamura *et al.* 2008; Sousa-Neves and Rosas 2010). This phenomenon is not limited to *Drosophila* as it is also found in other insects and plants (Lowry and Willis 2010). With the identification of genes that influence behavioural isolation more patterns will likely emerge, such as the type of genes that

produce female preference behaviour, the types of mutations that cause changes in the behaviour, and if these patterns are universal or genus-specific.

Unfortunately, the genetic basis of behavioural isolation remains unknown. Regions have been identified that are responsible for sexual isolation in various species of *Drosophila* (Moehring *et al.* 2004, 2006; Coyne 1992). However, without the use of *D. melanogaster* and plethora of genetic tools available in this species, the variants that contribute to species specific mating behaviour cannot be easily determined. The main hurdle is proper methodology. Mapping down to the gene level requires the tools that exist only in *D. melanogaster*. However, this species is not typically used in behavioural isolation research because the most widely employed technique (recombination mapping) requires fertile hybrids, and F_1 hybrids from both directions of crosses of *D. melanogaster* with its sibling species are sterile. Therefore, in order to identify a gene for interspecific mating preference in *D. melanogaster*, the mapping technique would have to measure F_1 hybrids rather than recombinant individuals. This is possible by using deficiency mapping.

Deficiency mapping employs a series of stocks available only in *D. melanogaster* that are fully diploid except for a small segment that has only one copy of the genes (hemizygous) due to a deletion on one of the homologous chromosomes. Each stock has a unique hemizygous region (deficiency). Multiple stocks with different deficiencies and varying size are available which allows for good coverage of the genome, and most stocks partially overlap with an adjacent deficiency. In an F_1 hybrid formed with parents from a *D. melanogaster* deficiency stock and the closely-related sibling species *D. simulans*, a homologous chromosome

will be inherited from each parent, resulting in loci that are represented by one allele from each species (heterozygous), except for in the deficiency region. Due to the deletion, the hybrid can only express the *D. simulans* alleles within this region in a hybrid background (Figure 2.1). Normally, female F_1 hybrids between these two species display *D. melanogaster*-like behaviour when interacting with a *D. melanogaster* male: F_1 hybrid females mate relatively readily, similar to *D. melanogaster* females, whereas *D. simulans* females demonstrate strong behavioural isolation against these males (Davis *et al.* 1996). However, if the deficiency contains a gene(s) for *D. simulans* female rejection of a *D. melanogaster* male, the *D. simulans* allele(s) are not masked due to the deficiency, and these females will have a lower mating propensity with *D. melanogaster* males compared to the hybrid female that did not inherit the deletion.

Previously, deficiency mapping was used to map the entire right arm of the third chromosome plus two regions overlapping genes previously identified through mutagenesis as influencing intraspecies female preference (Chapter 2 of this thesis). Five regions in total were found to be significant: one near the centromere (cytological region 81F3-82F7) one near the telomere (100A1-3Rt), and three within the inversion polymorphism between the two species (89B10-89E2; 91B2-91E4; and 93B7-93C5). Four of the five regions were fine-mapped by comparing the mating behaviour of hybrids created from deficiency lines that overlapping the previously identified areas. The other region (93B7-93C5) has not yet been fine-mapped due to time constraints. Fine mapping of the four regions not only decreased the total number of candidate genes from 481 to 262, but it also helped to

identify the first candidate gene for female preference, interspecific female preference, and behavioural isolation in *Drosophila*.

Materials and Methods

Stocks

Eighteen lines with deficiencies spanning four previously-identified significant regions (81F3-82F7; 89B7-89E7; 90F1-91F5; 99D3-3Rt) including two lines testing a candidate gene for behavioural isolation within these regions were obtained from the Bloomington Drosophila Stock Center (Bloomington, IN, USA; Table 3.1). All of the breakpoints were listed on the online database and initially provided by the donors. The chromosome that is affected by the deletion is termed the deficiency chromosome; the other contains serial inversions to reduce viability of recombinant offspring, contains a dominant visible marker, and is termed the balancer chromosome. The line is maintained *Df*/bal because both chromosomes are homozygous lethal. Two lines of wild-type *D. simulans* (FC, collected in Florida City; 216, collected in Scotland) were obtained from the Drosophila Species Stock Center (San Diego, CA, USA); wild-type *D. melanogaster* (BJS1) were collected in 2009 in London, ON, Canada by Dr. Brent Sinclair. Each line was kept in a standard 8-dram plastic vial and raised on approximately 7 ml of standard cornmeal and yeast medium (Bloomington Stock Center's standard medium recipe). All stocks were maintained in incubators with a LD 14:10 hour cycle, 25°C, and a relative humidity of 80%. One line was very weak (*Df*(3R)P9) and was transferred to a 21°C incubator, with all other components the same.

Table 3.1: Deficiency lines used to fine map four previously identified regions on 3R.

Deficiency	Region	Balancer	Marker	Notes
Df(3R)ME15	81F3-6;82F5-7	MKRS	Sb	BROAD MAP
Df(3R)ED5071	81F6;82E4	TM6c	Sb	completed
Df(3R)ED5138	82D5;82F8	TM6c	Sb	inviable
P{Mae-UAS.6.11)$_{FC}$	82F6	TM3	Sb	completed
P{Mae-UAS.6.11)$_{216}$	82F6	TM3	Sb	completed
Df(3R)BSC177	82F6;82F9	TM6b	Tb	completed
Df(3R)ED5156	82F8;83A4	TM6c	Sb	BROAD MAP
Df(3R)sbd105	88F9-89A1;B9-10	TM3	Ser	BROAD MAP
Df(3R)ED10639	89B7;89B18	TM6c	Sb	weak line
Df(3R)P115	89B7-8;89E7	TM1	Me	BROAD MAP
Df(3R)ED10642	89B17;89D5	TM6c	Sb	completed
Df(3R)P10	89C1-2;89E1-2	TM1	Me	weak line
Df(3R)P9	89E1;89E5	TM3	Sb	21C incubator
Df(3R)RD31	89E2;90D	In(3R)C	Sb	BROAD MAP
Df(3R)DG2	89E1-F4;91B1-B2	TM2	Ubx	BROAD MAP
Df(3R)BSC650	90C6;91A2	TM6c	Sb	weak line
Df(3R)Cha7	90F1-F4;91F5	TM6b	Tb	BROAD MAP
Df(3R)BSC509	91A3;91D5	TM6c	Sb	48-hour
Df(3R)ED5911	91C5;91F4	TM6c	Sb	completed
Df(3R)Dl-BX12	91F1-2;92D3-6	TM6b	Tb	BROAD MAP
Df(3R)L127	99B5-6;99F1	TM6	Ubx	BROAD MAP
Df(3R)B81	99D3;3Rt	TM3	Sb	BROAD MAP
Df(3R)tll-g	99F1-2;100B4-5	TM3	Sb, Ser	completed
Df(3R)A113	100A;100F	In(3R)C	Sb, Tb	weak line
Df(3R)BSC793	100B5;100C4	TM6c	Sb	completed
Df(3R)Exel6218	100B5;100C1	TM6b	Tb	weak line
Df(3R)ED6361	100C7;100E3	TM6c	Sb	completed
Df(3R)04661	100D2;100F5	TM3	Sb	weak line
Df(3R)ED50003	100E1;100E3	TM6c	Sb	completed

Completed = line was successfully completed; Inviable = *sim*/bal and/or *sim*/*Df* was inviable; Too weak = line was too weak to be maintained; 21 incubator = crosses were completed within a 21°C incubator; 48-hour = hybrids were paired with males for 48 hours before dissections; Broad map = line was previously tested and region may hold genes for interspecific female preference.

Crosses

To use deficiency mapping to locate the genes for interspecific female preference, each line was independently crossed to wild-type *D. melanogaster* and FC *D. simulans*. Female virgins of each stock were collected 0-8 hours after eclosion, separated under CO_2 anaesthesia, and transferred to new vials at low densities (1-20) to be housed for at least seven days to ensure virginity and reproductive maturity. To create the F_1 hybrid females, 10 female virgins from each deficiency stock and 20-25 0-7 day old FC *D. simulans* males were placed in a 8-dram plastic vial with 7 ml of medium (as described above) with the cotton pushed down to reduce the available space and thus increase interactions between the two species. Two types of heterospecific test hybrid females were produced from this cross: *sim*/bal and *sim/Df*. To control for effects of the balancer and deficiency chromosome on general mating behaviour, five female virgins from each deficiency stock and five 0-7 day aged *D. melanogaster* males were placed in an 8-dram plastic vial with 7 ml of medium (as described above). Two types of pure-species test females were produced from this cross: *mel*/bal and *mel/Df*. Each cross vial was transferred every 5 days until larvae were no longer produced.

In addition to the crosses described above, line P{Mae-UAS.6.11) was crossed to 216 *D. simulans* with the same method as the cross with FC *D. simulans*. Line 216 was used to investigate if different inbred lines of the *D. simulans* have the same genetic basis of interspecific female preference. Additional pure-species control hybrids were produced with the same method as listed above. Therefore four

additional types of test females were produced: *sim*216/bal; *sim*216/*Df*; mel/bal; and mel/*Df* the crosses are hereafter referred to as P{Mae-UAS.6.11)$_{216}$.

Mating assay

Test females were collected 0-8 hours after eclosion, separated based on the presence of the dominant marker (indicating the inheritance of the balancer chromosome) under light CO_2 anaesthesia, transferred to new vials of 1-10 flies, and housed for 5-7 days. Virgin wild-type *D. melanogaster* males were collected and housed the same way.

One test female was placed with one wild-type *D. melanogaster* male for 45 minutes in an 8 dram glass vial misted with water to increase humidity. Equal numbers of each type of test female (*sim/Df*, *sim*/bal, *mel/Df*, *mel*/bal) were observed simultaneously to control for environmental effects. The time the pair was placed in the vial, the time of courtship occurrence and the time of mating occurrence was recorded. From these measures, courtship latency (time to start of courtship), courtship occurrence (proportion of the total number of females that were courted by *D. melanogaster*), and copulation occurrence (proportion of the number of courted females that mated with *D. melanogaster*) were determined for each type of female in each line.

Interspecific hybrid female's mating behaviour with *D. melanogaster* males is reduced in comparison to full species hybrid females: only a small number of both *sim/Df* and *sim*/bal mated with *D. melanogaster* males within the 45-minute mating assay. Therefore, to increase the number of observed matings the length of the assay was increased and sperm presence was measured, hereafter referred to as the

'sperm assay.' After the 45-minute assay, *sim/Df* and *sim*/bal hybrid females and their *D. melanogaster* male partners were placed in a plastic vial with food (preparation of vial is described above). Equal numbers of each type of interspecific hybrid females were placed in a vial and the number of each type of female did not exceed five. After 24-28 hours, the female reproductive tract and spermathecae were dissected and scored for sperm presence under a light microscope. One line (*Df*(3RBSC506) had very low interspecific hybrid mating occurrence for both *sim/Df* and *sim*/bal hybrid females. Therefore, to further increase the number of observed matings, the length of time of the assay was increased to 48-50 hours and carried out as described above.

Data Analysis

For each line, I compared the courtship occurrence (courted or not courted) by the *D. melanogaster* male towards the female of each of the four genotypes (*sim*/bal, *sim/Df, mel*/bal, and *mel/Df*) with use of a G test ($\alpha<0.05$) in order to determine whether the level of male courtship differed between the genotypes. Also, an ANOVA ($\alpha<0.05$) was used for the continuous variables to compare courtship latency for the different types of females. This was used to test for the potential effect of male courtship behaviour that could bias the interpretation of the female mating behaviour. If all females were not courted equally, the difference could be misinterpreted as a reluctance to mate on the female's behalf.

The balancer chromosome in the deficiency lines have multiple inversion breakpoints that could potentially disrupted the normal functioning of the *D. melanogaster* gene. If this occurred, this disruption would act as another deficiency

uncovering the functioning *D. simulans* allele. If this allele functioned as a gene for interspecific female preference, these *sim*/bal females would be less likely to mate with *D. melanogaster* males compared to the *sim/Df* females of the same line. To test for this possibility, I grouped the *sim*/bal females of the different lines according to their balancer and compared the copulation occurrence (copulated or did not copulate) with the *D. melanogaster* male for the 45 minute assay and the 24 hour assay with a one-way ANOVA ($\alpha<0.05$). This was only possible for balancers where more than one line with the balancer was tested. In order to assess if the balancer's effect on mating was due to the breakpoints and not the genetic background of the lines, I grouped the *sim/Df* females of the different lines according to their balancer and compared the copulation occurrence (copulated or did not copulate) with the *D. melanogaster* male for the 45 minute assay and the 24 hour assay with a one-way ANOVA ($\alpha<0.05$).

In order to map genes for interspecific female preference, for each line the courtship and copulation occurrence data (binomial variables) were assessed with a G test ($\alpha<0.05$). However, if the calculated minimum expected value was lower than five, then a Fisher's Exact test ($\alpha<0.05$) was used. For a result to be significant, the data had to show a specific pattern: the *sim/Df* hybrids have a reduced amount of mating compared to the interspecies control (*sim*/bal) after those values are corrected for any effect of the deficiency and balancer genetic background (*mel/Df, mel*/bal). Although I had moderate sample sizes, behaviour may be influenced by multiple genes of small effect (Greenspan 1995). Therefore I also considered a line

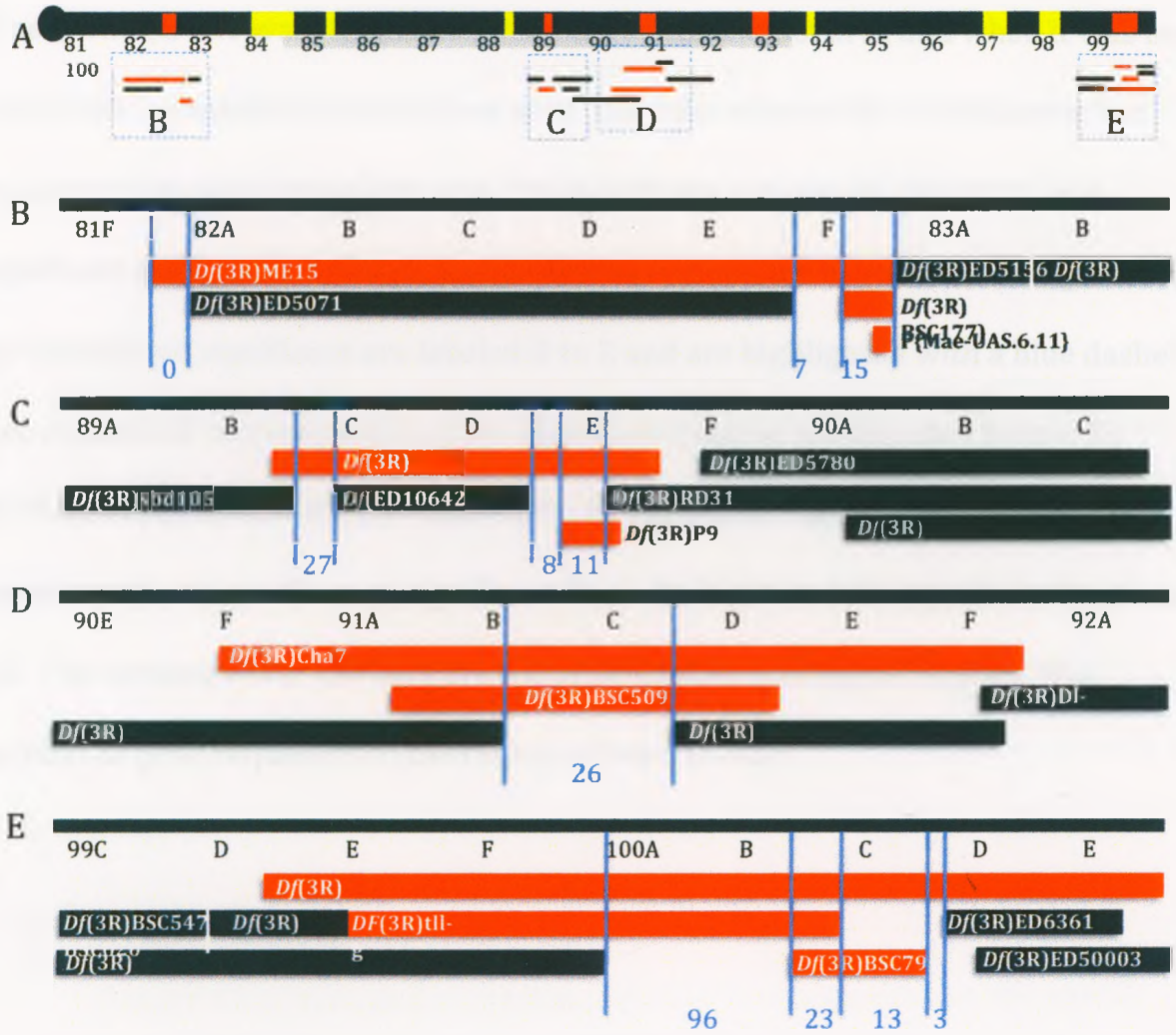
to be 'potentially significant' if $0.15>\alpha>0.05$ and in the expected direction. The courtship latency data (continuous variable) was assessed with an ANOVA ($\alpha<0.05$).

The previously identified significant lines (*Df*(3R)ME15, *Df*(3R)P115, *Df*(3R)Cha7, and *Df*(3R)B81) were compared to the overlapping fine-mapping lines (Figure 3.1) and the area considered to be significant was determined by interline comparison. When the fine-mapping line was not found to be significant, the portion of the deficiency of the original line that overlaps with the small non-significant fine mapping line was then considered to not have gene(s) for interspecific female preference. However, when the fine mapping line was found to be significant, it is most likely that portion of the deficiency of the original line that overlaps with the small significant fine-mapping line contains gene(s) for interspecific female preference, and therefore this area was highlighted as significant. Some of the deficiency lines had known breakpoints (base pair reported, Table 3.1) but some were estimated (estimated cytological location). If the fine-mapping line was of the latter, the minimal area of deficiency was considered for the non-significant lines and the maximum was considered for the significant lines as to not leave out a gene that may influence the trait.

Results

I used deficiency mapping to finer map 4 genomic regions along the right arm of the third chromosome to locate genes for interspecific female preference in *D. simulans*. I compared the mating behaviour of four types of females (*sim*/bal, *sim*/*Df*, *mel*/bal, and *mel*/*Df*) with *D. melanogaster* males. The null hypothesis is that the

Figure 3.1: Overlapping deficiencies spanning four regions of 3R used to fine map genes for interspecific female preference.



A representation of the deficiencies used to fine map interspecific mating preference in *D. simulans* females. A) The bar represents the right arm of the third chromosome. To the left, the circle represents the centromere and the blunt right end represents the telomere. The numbers under the bar are the approximate cytological regions and the gray bar represents the interspecific inversion polymorphism. Under the chromosome the deficiencies that covered 3R are represented as bars: the non-significant lines are black and the significant lines are red. The interline comparisons

were used to decrease genomic regions that were considered significant. Some of the lines have specific breakpoints listed (base pairs) and others have estimated breakpoints listed (cytological regions). In order to include all possible candidate genes, when a line was significant I considered the largest area and when it was not significant I considered the smallest area. The coloration of the chromosome bar represents the mapping of this arm: the significant regions are in red, the non-significant regions are in black, and un-tested regions are in yellow. The regions that are considered significant are labeled B to E and are highlighted with a blue dashed box. A detailed representation of the significant regions are depicted below: B) 81F3-82F2; C) 89B10-E2; D) 91B2-91F1; E) 99F2-3Rt. The deficiencies are represented as bars: the non-significant lines are black and the significant lines are red. The numbers over the bars are the approximate cytological regions. The number of gene sequences in the region is listed in blue.

region does not contain genes for behavioural isolation. If the *sim/Df* females mating propensity did not significantly differ from that of the *sim*/bal or this difference was also observed in the control females, the null hypothesis was not rejected. However, if the *sim/Df* females mated significantly less than the *sim*/bal and this difference was significantly different than that of the control females, the null hypothesis was rejected. The alternative hypothesis that there are gene(s) for interspecific female preference in the region would be accepted.

I attempted to complete 18 lines to fine map the four regions. Six lines were not completed because the interspecies hybrids were too weak for behaviour assays and one line because *sim*/bal females were inviable (Table 3.1).

Mating

For each line, I compared the copulation occurrence of the four types of hybrids with a *D. melanogaster* male for the 45 minute observational assay and the sperm assay using either a G test or a Fisher's Exact test ($\alpha<0.05$, or 'potentially significant' if $0.15>\alpha>0.05$).

For the 45 minute assay, one line (*Df*(3R)P9) was significant: the *sim/Df* females mated significantly less with *D. melanogaster* males compared to the other 3 types of females (Fisher's Exact test p=0.04), and one line (*Df*(3R)tll-g) was potentially significant: the *sim/Df* females had reduced mating with these males (G=2.79 p=0.09; Table 3.2).

For the 24 hour sperm assay, five lines had significantly less mating between the *sim/Df* females and *D. melanogaster* males compared to the other 3 types of females: *Df*(3R)P9 (G=6.72, p=0.01), *Df*(3R)BSC509 (G=17.75 p=0.00003), *Df*(3R)tll-

Table 3.2: Mating occurrence of the four types of females with *D. melanogaster* males.

		mel/bal		*mel*/*Df*		*sim*/bal		*sim*/*Df*			Sperm Assay			
***Df*(3R)**	**N**	**Crt**	**Cop**	**Crt**	**Cop**	**Crt**	**Cop**	**Crt**	**Cop**	**G test/Fishers**	**N**	***sim*/bal**	***sim*/*Df***	**G test/Fisher's**
ED5071	15	9	8	12	9	13	1	13	7	G=5.33 p=0.02 p=0.05	15	9	11	G=0.02 p=0.89
P{Mae-UAS.6.11}	20	16	13	15	11	13	4	12	1	G=1.86 p=0.17 p=0.20	51	32	12	G=4.91 p=0.03
P{Mae-UAS.6.11}	28	21	13	26	19	22	3	19	2	G=0.26 p=0.61 p=1	38	19	16	G=0.59 p=0.44
BSC177	32	27	21	30	24	31	1	29	0	G=0.91 p=0.34 p=1	72	17	5	G=6.14 p=0.01
ED10642	20	19	15	19	14	17	3	16	1	G=0.85 p=0.35 p=0.61	20	15	17	G=0.15 p=0.69
P9	31	23	20	24	12	30	16	27	1	G=4.35 p=0.04 p=0.04	30	22	6	G=6.72 p=0.01
BSC509	20	14	10	18	15	15	0	18	0	G=0 p=1 p=1	20	20	1	G=17.75 p = 0.000025
ED5911	20	17	17	19	11	17	0	20	2	G=3.81 p=0.05 p=0.15	41	20	12	G=0.03 p=0.87
tll-g	40	35	29	36	24	31	11	36	3	G=2.79 p=0.09	32	28	8	G=4.68 p=0.03
BSC793	20	14	11	14	10	13	1	16	0	G=1.92 p=0.16 p=1	20	8	0	G=8.30 p=0.004 p=0.03
ED6361	20	15	9	15	12	16	2	19	5	G=0.29 p=0.59 p=0.68	0	NA	NA	
ED50003	20	18	11	17	13	12	1	16	6	G=2.13 p=0.14 p=0.11	20	13	8	G=1.17 p=0.28

g (G=4.68 p=0.03), *Df*(3R)BSC177 (G=6.14, p=0.01), and *Df*(3R) BSC793 (Fisher's Exact test p=0.03).

Based on the significant and non-significant results of fine-mapping, the cytological regions that most likely contain gene(s) for interspecific female preference are: 81F3-6, 82F6-F7, 89B10-16, 89D6-E2, 91B2-C4, 99F2-100C6, and 100F5-3Rt, as well as 93B7-C5 which was not fine mapped (Figure 3.1).

For line that disrupted *katanin-60* gene with the FC background, the females did not significantly differ in their mating behaviour with *D. melanogaster* males during the 45-minute assay (Fisher's Exact test p=0.20), however, for the 24 hour sperm assay, there was significantly less mating between the *sim/Df* females with *D. melanogaster* males compared to the other 3 types of females (G=4.90, p=0.03). For line that disrupted *katanin-60* gene with the 216 background, the females did not significantly differ in their mating behaviour with *D. melanogaster* during the 45-minute assay (Fisher's Exact test p=1), nor for the 24 hour sperm assay (G=0.59, p=0.44).

Male Courtship Behaviour

For each line, I compared the courtship occurrence (courted or not courted) by the *D. melanogaster* male towards the female of each of the four genotypes (*sim*/bal, *sim/Df*, *mel*/bal, and *mel/Df*) with use of a G test ($\alpha<0.05$) in order to determine whether the level of male courtship differed between the genotypes. No lines produced a significant value (Table 3.3). For each line, I also compared the courtship latency (time until courtship began) by the *D. melanogaster* male towards the female of each of the four genotypes with use of a two-way ANOVA ($\alpha< 0.05$). No

Table 3.3: *D. melanogaster* courtship occurrence toward the four types of females.

Df(3R)	*mel*/bal	*mel*/*Df*	*sim*/bal	*sim*/*Df*	G test
ED5071	9	12	13	13	G=0.238, p=0.625
P{Mae-UAS.6.11)$_{FC}$	16	15	13	12	G=0.000831, p=0.977
P{Mae-UAS.6.11)$_{216}$	21	26	22	19	G=0.707, p=0.4
BSC177	27	30	31	29	G=0.216, p=0.642
ED10642	19	19	17	16	G=0.016, p=0.899
P9	23	24	30	27	G=0.141, p=0.708
BSC509	14	18	15	18	G=0.019, p=0.89
ED5911	17	19	17	20	G=0.012, p=0.913
Tll-g	35	36	31	36	G=0.127, p=0.722
BSC793	14	14	13	16	G=0.153, p=0.696
ED6361	15	15	16	19	G=0.119, p=0.73
ED50003	18	17	12	16	G=0.459, p=0.498

lines produced a significant value for species or genotype main effect (Table 3.4). One line had a significant interaction effect due to the rate at which the deficiencies were courted (*Df*(3R)ED5911: F(1,69)=4.11, p=0.05).

Effect of balancer

In the 11 lines I successfully used, there were three different balancers. I was not able to test the TM6B balancer because it was only present in line *Df*(3R)BSC177. Hybrid *sim*/bal females with the balancer TM3 had significantly higher copulation occurrence with *D. melanogaster* males than hybrid *sim*/bal females with the balancer TM6C during the 45-minute observational assay (p=0.05). However, this relationship broke down over time and the two groups of females did not significantly differ in copulation occurrence after the 24 hour sperm assay (p=0.60; Table 3.5). Hybrid *sim*/*Df* females with the balancer TM3 did not significantly differ in copulation occurrence, however, the average copulation occurrence of *sim*/*Df* females from TM3 balancer lines was greater than that of the TM6C *sim*/*Df* females (non-significant trend; Table 3.6).

Discussion

I fine mapped four of the five regions on the right arm of the third chromosome previously identified to be important for isolating two species of *Drosophila*. Furthermore, I investigated the first candidate gene for female interspecific female preference.

Table 3.4: *D. melanogaster* courtship latency toward the four types of females.

Df(3R)	*mel*/bal mean/SD	*mel*/*Df* mean/SD	*sim*/bal mean/SD	*sim*/*Df* mean/SD	Species	Genotype	Interaction
ED5071	21.6/11.90	13.5/12.20	14.8/11.80	10.8/11.50	F(1,43)=1.84 p=.18	F(1,43)=3.04 p=.09	F(1,43)=0.34 p=.56
P{Mea}$_{FC}$	15.09/12.0	11.02/9.89	7.05/3.72	11.97/11.00	F(1,52)=1.78 p=.19	F(1,52)=0.03 p=.87	F(1,52)=2.89 p=.09
P{Mea}$_{216}$	18.56/12.8	13.71/9.55	11.80/9.46	15.02/13.99	F(1,84)=1.24 p=.27	F(1,84)=0.11 p=.74	F(1,84)=2.72 p=.10
BSC177	12.23/8.47	15.17/11.4	10.70/6.56	12.21/10.01	F(1,113)=1.72 p=.19	F(1,113)=1.7 p=.20	F(1,113)=0.18 p=.68
ED10642	14.46/14.9	14.41/13.6	14.81/12.2	14.24/10.16	F(1,67)=0.001 p=.98	F(1,67)=0.01 p=.92	F(1,67)=0.01 p=.93
P9	18.07/13.6	14.69/11.4	11.47/9.34	13.59/10.75	F(1,100)=3.02 p=.09	F(1,100)=0.08 p=.76	F(1,100)=0.08 p=.77
BSC509	13.85/8.92	11.79/8.41	12.85/13.1	9.38/8.95	F(1,61)=0.48 p=.49	F(1,61)=1.25 p=.27	F(1,61)=0.08 p=.78
ED5911	7.37/7.29	12.82/10.2	11.22/12.0	7.98/6.12	F(1,69)=0.05 p=.82	F(1,69)=0.26 p=.61	F(1,69)=4.11 p=.05
tll-g	12.63/10.6	12.65/9.91	11.82/11.1	11.79/9.72	F(1,134)=0.23 p=.63	F(1,134)=0.01 p=.99	F(1,134)=0 p=.99
BSC793	17.60/11.5	17.16/12.0	20.56/12.1	11.15/11.92	F(1,53)=0.23 p=.63	F(1,53)=2.43 p=.12	F(1,53)=2.02 p=.16
ED6361	14.98/11.6	12.41/12.2	12.79/8.14	11.11/8.06	F(1,61)=0.49 p=.49	F(1,61)=0.72 p=.40	F(1,61)=0.03 p=.86
ED50003	14.99/8.19	17.4/12.93	13.17/9.18	13.48/11.17	F(1,59)=1.13 p=.29	F(1,59(=0.25 p=.62	F(1,59)=0.15 p=.60

Table 3.5: The effect of the balancer chromosome on female mating behaviour with *D. melanogaster* males.

Df(3R)	Balancer	Court	Cop	Prop mated	Balancer mean	Balancer SD	N	sperm	Prop	Balancer mean	Balancer SD
P{Mea}$_{FC}$	TM3	13	4	0.3077	0.3331	0.1632	51	32	0.6275	0.6839	0.1591
P{Mea}$_{216}$	TM3	22	3	0.1364			38	19	0.5		
P9	TM3	30	16	0.5333			30	22	0.7333		
tll-g	TM3	31	11	0.3548			32	28	0.875		
BSC177	TM6b	31	1	0.0323	NA	NA	72	17	0.2361	NA	NA
ED10642	TM6c	17	3	0.1765	0.0770	0.0634	20	15	0.75	0.6480	0.2116
ED5071	TM6c	13	1	0.0769			15	9	0.6		
BSC509	TM6c	15	0	0			20	20	1		
ED5911	TM6c	17	0	0			41	20	0.4878		
BSC793	TM6c	13	1	0.0769			20	8	0.4		
ED50003	TM6c	12	1	0.0833			20	13	0.65		
ED6361	TM6c	16	2	0.125			0	NA	NA		

Table 3.6: The effect of the genetic background of the deficiency chromosome on female mating behaviour with *D. melanogaster* males.

Df(3R)	Balancer	Court	Cop	Prop mated	Balancer mean	Balancer SD	N	sperm	Prop	Balancer mean	Balancer SD
P{Mea}$_{FC}$	TM3	12	1	0.0833	0.0772	0.0287	51	12	0.2353	0.2766	0.0986
P{Mea}$_{216}$	TM3	19	2	0.1053			38	16	0.4211		
P9	TM3	27	1	0.0370			30	6	0.2		
tll-g	TM3	36	3	0.0833			32	8	0.25		
BSC127	TM6b	29	0	0	NA	NA	72	5	0.0694	NA	NA
ED10642	TM6c	16	1	0.0625	0.1913	0.2072	20	17	0.85	0.3877	0.3483
ED5071	TM6c	13	7	0.5385			15	11	0.7333		
BSC509	TM6c	18	0	0			20	1	0.05		
ED5911	TM6c	20	2	0.1			41	12	0.2927		
BSC793	TM6c	16	0	0			20	0	0		
ED50003	TM6c	16	6	0.375			20	8	0.4		
ED6361	TM6c	19	5	0.2632			0	NA	NA		

Male Courtship Behaviour

For each line, *D. melanogaster* courted the hybrid females as often as the controls, and the deficiencies as often as the balancers. Furthermore, courtship latency did not differ for most lines. This further supports the notion that male *D. melanogaster* courted the hybrid females as rapidly as the controls, and the deficiencies as rapidly as the balancers. One line had a significant interaction effect due to the rate at which the deficiencies were courted but in the opposite direction that would cause it to bias the interpretation of the female receptivity results (*mel*/*Df* were courted more slowly than *sim*/*Df*).

Effect of the balancer

The effect of the balancer on mating behaviour was examined in order to ensure that the breakpoints of the inversions along the chromosome did not disrupt a gene for interspecific female mating behaviour gene. This would be obvious as the mating behaviour of *sim*/bal females when paired with *D. melanogaster* males would be very low compared to *sim*/*Df* females and the controls. Although the TM6C *sim*/bal female mating occurrence was significantly lower than the TM3 *sim*/bal females, it is unlikely that the TM6C balancer's breakpoints disrupted a gene for interspecific female preference as three of the six significant lines had this balancer. If the breakpoints disrupted gene(s) for interspecific preference, the *sim*/bal mating occurrence would be too low to produce a significant relationship.

While investigating the effect of the balancer it became apparent that interspecific hybrids with the TM3 balancer had increased mating occurrence with *D. melanogaster* males compared to hybrids with other balancer types. This could

bias the interpretation of female interspecific mating behaviour by creating a large difference between the interspecific hybrids which could be incorrectly seen as a significant result. In other words, *sim*/bal females had relatively high mating propensities with *D. melanogaster* males (compared to *sim*/*Df*) due to the balancer chromosome. However, it could be incorrectly assumed that the relatively low mating propensity of *sim*/*Df* females was due to the deficiency uncovering the *D. simulans* genome. Although three of the four lines with the TM3 balancer are significant, it is unreasonable to assume that in increase in mating behaviour occurrence of *sim*/bal females was driving the relationship. The difference in the behaviour persisted during a 24-hour period and was observed with the sperm assay which is after the relative increase in mating propensity of the *sim*/bal females broke down.

An alternative explanation is that the *D. melanogaster* background in which the TM3 balancer was produced may be responsible for the increased mating behaviour. If this was true, both interspecific females (*sim*/bal and *sim*/*Df*) would show an increase in mating occurrence compared to the hybrid females in the other lines. To investigate the background genetic effect of TM3 balancer lines, I grouped the *sim*/*Df* females of the different lines according to their balancer and compared the copulation occurrence with the *D. melanogaster* male for the 45-minute assay and the 24-hour assay with a t-test ($\alpha<0.05$; Table 3.4). The mating behaviour of the *sim*/*Df* females of the different balancer types did not significantly differ TM3 females had a non-significant trend of greater mating propensity.

For the best comparison, the significant lines should be removed, as these lines would decrease the mating average occurrence due to the unmasking of the *D. simulans* genome rather than due to the genetic background of the *D. melanogaster* line. However, they were not removed because TM3 would not have been able to be compared since three out of the four lines were found to be significant due to the effect of the deficiency. This may have contributed to the non-significant trend that was found. In the future, more lines should be tested and only those that did not have *sim*/*Df* females with significantly lower mating behaviour be used to the analysis. The cause of the increase mating behaviour of female with the TM3 balancer compared to the others remains unknown, however, I feel confident that it is most likely due to background genetic effects. This effect was controlled for as I compared the mating behaviour of *sim*/bal to *sim*/*Df* females of the same line.

Mating

Both lines that were considered significant or potentially significant during the 45-minute mating assay were found to be significant for the sperm assay. Lines *Df*(3R)BSC509, *Df*(3R)BSC177, and *Df*(3R)BSC793 were not found to be significant or potentially significant during the 45-minute assay but found to be significant during the sperm assay. All three of these lines had extremely low overall mating of the interspecies hybrids (both *sim*/*Df* and *sim*/bal) within the 45 minute assay, and thus the effect of the deficiency was likely only able to be detected with a longer assay period that increased the overall number of matings (primarily for the *sim*/bal genotype in these cases).

I fine mapped four of the five regions previously identified to influence the rejection of *D. melanogaster* males by *D. simulans* females (Figure 3.1). The regions were originally identified through broad deficiency mapping of the entire right arm of the third chromosome (Chapter 2 of this thesis). This part of the genome was chosen due to the importance of this region in the behavioural isolation of other *Drosophila* species (Noor *et al.* 2001; Moehring *et al.* 2006; Sawamura *et al.* 2007, 2008) and the presence of an interspecific inversion polymorphism, which has been shown to be important in species isolation in *Drosophila* (Noor *et al.* 2001; Nanda and Singh 2001; Sawamura *et al.* 2008; Sousa-Neves and Rosas 2010), other animals (Ayala *et al.* 2011), and plants (Lowry and Willis 2010).

I further fine mapped the four regions through fine deficiency mapping using smaller deficiencies that overlapped the original regions (Figure 3.1). Based on the significant and non-significant results of fine mapping, the regions that most likely contain gene(s) for interspecific female preference are: 81F3-6; 82E5-F7; 89B10-16; 89D6-E2; 91B2-C4; 99F2-100C6, and 100F5-3Rt, as well as 93B7-C5 which was not fine mapped in this study (Figure 2). Multiple gene sequences are in the small regions: 27 candidate genes in 89B10-16; 19 candidate genes in 89D6-E2; 26 candidate genes in 91B2-C4; 33 candidate genes in 93B7-C5; 133 candidate genes in 99F2-100C6; and 22 candidate genes in 82E5-F7 including the first candidate gene for interspecific female preference *katanin-60*. No gene sequences are located in 81F3-6 and 100F5-3Rt regions. In total, 260 genes (Appendix A) have been identified to possibly influence interspecific preference, with the most compelling evidence for *katanin-60*.

Both the mutant that only disrupted *katanin-60* as well as the deficiency that removed the *katanin-60* promoter acted to uncover *D. simulans* FC genome in the interspecies hybrid and provided evidence that this gene may be contributing to the behavioural isolation seen between these species. However, when the deficiency that contained *katanin-60* was used to uncover *D. simulans* 216 genome, it did not significantly affect interspecies female preference. This may be due to the inbred nature of the lab lines which capture just a snapshot of the genetic diversity from the populations in which they were derived from. Therefore, it is possible that this variant may act to isolate *D. simulans* FC and *D. melanogaster* but it may not act in *D. simulans* 216, and therefore, may not be species wide. Similarly, Ortiz-Barrientos et al. (2004) identified two regions on the X chromosome for behavioural isolation between *D. psuedoobscura* and *D. persimilis*. In an effort to replicate the findings, an independent QTL study was done, but the same regions were not found (Barnell and Noor 2008). Likewise, the same behaviour that was mapped in this study (rejection of *D. melanogaster* males by *D. simulans* females) was also mapped broadly to the chromosome in other studies, with conflicting results. Although two studies found an effect of the X chromosome, one also found an effect of the 3rd chromosome (Uenoyama and Inoue 1995), while the other identified the left arm of the 2nd chromosome (Carracedo *et al.* 1998). *D. simulans* spans a large geographic area. It is possible that, although the populations of the same species are able to successfully mate, the genomes have diverged to create different genetic variants. These two lines come from different locations (Florida and Scotland), which have different climates, and possibly different selective pressures. Although *katanin-60* may not

contribute to behavioural isolation for the 216 line, it is still a viable candidate gene for behavioural isolation as it does act to reduce interspecific mating in another population.

Katanin-60 was first discovered as an ATP-dependent microtubule severing gene that is responsible for the organization of cell infrastructure in sea urchin eggs (McNally and Vale 1993). Two homologs have been identified, p60 and p80, that have been found in many species from plants to humans (Roll and McNally 2010). *Katanin-60* is expressed in the central nervous system of *Drosophila*, including the mushroom body, and functions to regulate axonal outgrowth and branching (Ahmad and Bass 1999; Zhang *et al.* 2011). The mushroom body is important for the integration of many sensory systems in *Drosophila* (Davis 2011), especially olfaction (de Belle and Heisenberg 1994), and has been linked to sexual behaviour (O'dell *et al.* 1995; Balakireva *et al.* 1998). *Drosophila* male courtship is complex with multiple signals, which can each be broken down into their own components. For example, the male cuticular hydrocarbon (CHC) blend, used as a pheromone, is species specific and is a combination of at least 20 or more hydrocarbons; the male courtship song is a combination of pulse and sine song and these vary in a species-specific manner (Ferveur 2010). Therefore, it is possible that the mushroom body integrates the various signals from a courting male: if the signals are correct it causes the female to be receptive, but if the different species' alleles gave rise to different neuronal growth and branching in the mushroom body, it is possible that it could cause females of different species to be stimulated only by the species-specific combination of male signals. However, the role of *katanin-60* in female receptivity is

purely speculative at this stage as transgenics are required to confirm that *katanin-60* is a gene for interspecific female preference.

Unlike microarray studies, where the identified candidate genes are all likely to play some role in the measured trait, I presume only one or very few candidate genes within each of the regions contributes to female preference behaviour. Thus, a meta-analysis of all of the candidate genes within each region is highly unlikely to reveal a significant trend or pathway. However, I can elucidate potential trends or likely candidates within these regions based on known gene function. Of the remaining genes in the fine-mapped regions, there are a number of potential candidates for behavioural isolation. The most obvious to highlight as a candidate gene are those involved in the olfactory system, which is involved in the sensation of pheromones. *Drosophila* use CHC profiles to communicate their species and sexual identity, and these profiles have been shown to be important in mating behaviour (Ferveur 2005), are used to determine suitable mates (Billiter *et al.* 2009), and contribute to behavioural isolation between species (Coyne *et al.* 1994). Differences in an olfactory gene and gene product could cause females to be more sensitive to different CHC profiles and lead to a difference in mate choice. Multiple genes in the regions identified through deficiency mapping are involved in olfaction such as genes expressed in the olfactory system (*gish*), involved in odorant binding (*CG31189*, *CG7079*, and *Obp93a*), and those specifically found to be responsible for pheromone sensation (*Snmp1*). CHCs are nonvolatile chemicals, and therefore, it is possible that the sense of taste is also involved. Taste has previously been implicated in mating behaviour as important for continual male courtship (Robertson 1982).

Therefore, genes that are involved with taste (*Mvl*) could be involved with behavioural isolation as perception of food odours and mating behaviour have been previously linked (Grosjean *et al.* 2011). Finally, it is possible that the genetic basis for behavioural isolation occurs earlier on in development of the sensory organs (*Rab11*, *spdo* and *wts*).

Sex determinism pathway genes such as *fruitless* and *doublesex* have been found to be important for normal male and female mating which is most likely due to the sexual identity of sex-specific neurons (for a review Ferveur 2010). Although one gene (*dmrt98B*) was identified as a candidate gene, it is unlikely that this gene influences species typical mating behaviour as it does for sex-specific behaviour).

The rejection of a heterospecific male by a female could be due to different timing or location of gene expression, or due to a slight change in the protein product by alternative splicing. It has been shown that mutations in transcription factors involved in the sex-determination pathway, such as *fruitless*, can influence male mating behaviour (Hall 1978), and alternate splicing in the gene product can influence the presence of male- and female-typical behaviour (for review, see Ferveur 2010). Furthermore, one of the genes identified through mutagenesis as to influence intraspecific female receptivity, *dissatisfaction,* is a transcription factor (Finley *et al.* 1998). Species-specific changes in gene products, increases/decreases in expression, or alternative gene products due to splicing could have downstream effects that influence any number of pathways that could subsequently influence interspecific female receptivity. All of my candidate regions contain genes that are involved with various aspects of gene expression: regulation of transcription (*abd-A*,

tre-1, *tre-2*, *tre-3*, *dmrt93B*, *CG7056*, and *Sox100B*), gene silencing (*corto*), translation (*tRNA:F2:89BC*, *tRNA:V4:89BC*, *tRNA:CR31282*, *tRNA:CR31497*, *mRpS33*, *tRNA:CR31228*, *mRpS18C*, *CG1340*, *mRpL32*, and *CG11334*), alternative splicing (*snRNA:U4atac:82E*, and *nonA-l*), RNA processing (*Cpsf73*, and *CG7009*), and methylation (*CG14906*).

It has been predicted that a gene for behaviour would most likely be expressed in the brain because all three genes for intraspecific female receptivity found through mutagenesis (*spinster*, *dissatisfaction*, and *chaste*) are expressed in the central nervous system (Suzuki *et al.* 1997; Finley *et al.* 1998; Juni and Yamamoto 2009). It is possible that the differences in female receptivity behaviour seen between species may be due to a difference in how the brain and nervous system function. Although each female may be able to receive various signals produced by the male (pheromone, auditory, and tactile signals) she may interpret and respond to them differently due to the species-specific neuronal circuitry. This is consistent with the finding that females are stimulated less by heterospecific courtship song than conspecific song (Immonen and Ritchie 2011). Furthermore, it has been shown that neurotransmitters can influence male courtship (Zhang and Odenwald 1995) and female interest (Becnel *et al.* 2011). Finally, many of the genes important in assortative mating in the Z strains of *D. melanogaster* are highly expressed in the central nervous system (Bailey *et al.* 2011). Therefore genes involved in central or peripheral nervous system development (*ss*, *Manf*, *fray*, *RhoGAP93B*, *sec15*, *Ptx1*, *Aph-4*, and *zfh1*), neurotransmitter activity (*5-HT7*), or synaptic transmission (*Cdase*) could influence behavioural isolation.

Just as male courtship behaviour is stereotypical, so too is female receptivity. When a female becomes more receptive to a male she will slow down her locomotor behaviour, orient herself to the male, present her abdomen, and spread her wings to allow the male to copulate (Hall 1994; Greenspan 1995; Greenspan and Ferveur 2000; Griffith and Ejima 2009). Therefore genes that are involved in locomotion (*Dhc93AB*, and *dj-1beta*) could have species-specific variants causing females to either move differently or slow their behaviour for different species-specific signals.

Previous QTL studies identifying regions for behavioural isolation and microarray studies identifying natural variation in gene expression and mating behaviour have not identified regions that contain genes found through mutagenesis as being important for normal intraspecific mating behaviour (Carney 2007; Ruedi and Hughes 2009; Moehring and Mackay 2004; Gleason *et al.* 2005; Mackay *et al.* 2005). Therefore, although the genes are important in the construction of normal behaviour, they have not been found to contribute to variation within or even between species, and therefore do not likely contribute to behavioural isolation. Although females of two different species may be paying attention to the same male traits (Immonen and Ritchie 2011), the genetic architecture of within and between species mate choice may differ (Blows and Allan 1998). Nonetheless, a gene that has been found to influence male courtship behaviour (*5-HT7*) was present in one of the candidate regions. Furthermore, the gene *nonA* has been found in influence male courtship song (Kulkarni and Hall 1987), and the related gene *nonA-l* (which has similar properties and functions) was also present in one of the regions. Although it has not been confirmed by experimental research, it is possible that a gene

responsible for the mating behaviour within a species could tolerate mutations, which creates variation. This mutation would most likely not stop the behaviour but simply slightly change it. For example, in female mating behaviour, mutations to genes responsible for intraspecific female receptivity would be tolerated if it altered the specific level of the male trait that most excited the females (e.g. IPI) but still allowed for the species typical signal to excite the female enough for successful copulation. During the process of speciation, different variants may become fixed in different populations, and that variation could lead to the behavioural isolation of two species. For example, subtle changes in genes that influence the timing of general increased activity or timing of increased mating propensity (such as *5-HT7*, which has been found to be important in activity levels and circadian rhythm; Becnel *et al.* 2011) could lead to behavioural isolation over evolutionary time.

Conclusions

This study identified small regions that may contain gene(s) contributing to behavioural isolation between *D. simulans* and *D. melanogaster*. What sets this research apart from previous studies that also identified similar regions is the use of the model organism *D. melanogaster*. This allows for the fine-mapping of these regions down to the gene level due to the multitude of genetic tools available in this species. Additional small deficiency lines, and single gene mutation lines that perform as a deficiency line for a single gene, can be tested to reduce the number of candidate genes and identify the variants that contribute to behavioural isolation. This was accomplished with *katanin-60*, the first behavioural isolation candidate gene, and the remaining regions and genes can be tested following the same

methodology. The eventual confirmation of *katanin-60* as a gene contributing to behavioural isolation, and the identification of additional individual genes that influence female preference between these two species of *Drosophila*, will contribute to the knowledge of interspecific female preference, behavioural isolation, and the process of speciation.

References

Ahmad, F. J., W. Yu, F. J. McNally, and P. W. Baas. 1999. An essential role for *katanin* in severing microtubules in the neuron. J. Cell. Biol. 145:305-315.

Ayala, D., M. C. Fontaine, A. Cohuet, D. Fontenille, R. Vitalis, and F. Simard F. 2011. Chromosomal inversions, natural selection and adaptation in the malaris vector *Anopheles funestus*. Mol. Biol. Evol. 28:745-758.

Bailey, R. I., P. Innocenti, E. H. Morrow, U. Friberg, A. Qvernstrom. 2011. Female *Drosophila melanogaster* gene expression and mate choice: the X chromosome harbours candidate genes underlying sexual isolation. PLoS ONE 6:e17358.

Barnell, C. V, and M. A. Noor. 2008. Failure to replicate mate preference QTLs across multiple strains of *Drosophila psuedoobscura*. J. Hered. 99:653-656.

Bayes, J. J., and H. S. Malik. 2009. Altered heterochromatin binding by a hybrid sterility protein in *Drosophila* sibling species. Science 326:1538-1541.

Becnel, J., O. Johnson, J. Luo, D. R. Nassel, C. D. Nichols. 2011. The serotonin 5-HT7Dro receptor is expressed in the brain of *Drosophila*, and is essential for normal courtship and mating. PLoS ONE 6:e20800.

Billiter, 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.

Blows, M. W., and R. A. Allan. 1998. Levels of mate recognition within and between two *Drosophila* species and their hybrids. Am. Nat. 152:826-837.

Carney, G. E. 2007. A rapid genome-wide response to *Drosophila melanogaster* social interactions. BMC Genomics 8:288.

Carracedo, M. C., R. Pineiro, and P. Casares. 1995. Chromosomal substitution analysis of receptivity and sexual isolation in *Drosophila melanogaster* females. Heredity 75:541-546.

Carracedo, M. C., A. Suarez, A. Asenjo, and P. Casares. 1998. Genetics of hybridization between *Drosophila simulans* females and *D. melanogaster* males. Heredity 80:17-24.

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

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., and H. A. Orr. 1997. Patterns of speciation in *Drosophila* revisited. Evolution 51: 295-303.

Davis, A. W., J. Roote, T. Morley, K. Sawamura, S. Herrmann, and M. Ashburner. 1996. Rescue of hybrid sterility in crosses between *D. melanogaster* and *D. simulans*. Nature 380:157-159.

Davis, R. L. 2011. Traces of *Drosophila* memeory. Neuron 14:8-19.

de Belle, J. S., and M. Heisenberg. 1994. Associative odor learning in *Drosophila* abolished by chemical ablation of mushroom bodies. Science 263:692-695.

Doi, M., M. Matsuda, M. Tomaru, H. Matsubayashi, and Y. Oguma. 2001. A locus for female discrimination behavior causing sexual isolation in *Drosophila*. Proc. Natl. Acad. Sci. USA 98:6714-6719.

Ferveur, J. F. 2005. Cuticular hydrocarbons: their evolution and roles in *Drosophila* pheromonal communication. Behav. Genet. 35:279-295.

Ferveur, J. F. 2010. *Drosophila* female courtship and mating behaviors: sensory signals, genes, neural structures and evolution. Curr. Opin. Neurobiol. 20:764-769.

Finley, K. D., P. T. Edeen, M. Foss, E. Gross, N. Ghbeish, P. H. Palmer, B. J. Taylor, and M. McKeown. 1998. *Dissatisfaction* encodes a tailless-like nuclear receptor expressed in a subset of CNS neurons controlling *Drosophila* sexual behavior. Neuron 21:1363-1374.

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.

Greenspan, R. J. 1995. Understanding the genetic construction of behavior. Sci. Am. 271:74-79.

Greenspan, R. J., and J. M. Ferveur. 2000. Courtship in *Drosophila*. Annu. Rev. Genet. 34:205-232.

Griffith, L. C., and A. Ejima. 2009. Multimodal sensory integration of courtship stimulating cues in *Drosophila melanogaster*. Ann. N. Y. Acad. Sci. 1179:394-398.

Grosjean, Y., R. Rytz, J. P. Farine, L. Abuin, J. Cortot, G. S. Jefferis, and R. Benton. 2011. An olfactory receptor for food-derived odours promotes male courtship in *Drosophila*. Nature: 478:236-240.

Hall, J.C. 1978. Courtship among males due to a male-sterile mutation in *Drosophila melanogaster*. Behav. Genet. 8:125-141.

Hall, J. C. 1994. The mating fly. Science 264:1702-1714.

Immonen, E. and M. G. Ritchie. 2011. The genomic response to courtship song stimulation in female *Drosophila melanogaster*. Proc. Biol. Sci. Epub ahead of print.

Jamart, J. A., M. C. Carracedo, and P. Casares. 1993. Sexual isolation between *Drosophila melanogaster* females and *D. simulans* males. Male mating propensities versus success in hybridization. Experientia 49:596-598.

Juni, N., and D. Yamamoto. Genetic analysis of *chaste*, a new mutation of *Drosophila melanogaster* characterized by extremely low female sexual receptivity. 2009. J. Neurogenet. 23:329-340.

Kulkarni, S. J., and J. C. Hall. 1987. Behavioural and cytogenetic analysis of the cacophony courtship song mutant and interacting genetic variants in *Drosophila melanogaster*. Genetics 115:461-475.

Lee, H. Y., J. Y. Chou, L. Cheong, N. H. Chang, S. Y. Yang., and J. Y. Leu. 2008. Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. Cell 135:1065-1073.

Lowry, D. B., and J. H. Willis. 2010. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. PLoS ONE 8:e1000500.

Mackay, T. F., S. L. Heinsohn, R. F. Lynman, A. J. Moehring, T. J. Morgan, and S. M. Rollmann. 2005. Genetics and Genomics of *Drosophila* mating behavior. Prac. Natl. Acad. Sci. USA 102:6622-6629.

Masly, J. P., C. D. Jones, M. A. Noor, J. Locke, and H. A. Orr. 2006. Gene transposition as a cause of hybrid sterility in *Drosophila*. Science 8:1448-1450.

McNally, F. J., and R. D. Vale. 1993. Identification of Katanin, an ATPase that severs and disassembles stable microtubules. Cell 75:419-429.

Mihola, O., Z. Trachtulec, C. Vlcek, J. C. Schimenti, and J. Forejt. 2009. A mouse speciation gene encodes a meiotic histone H# methyltransferase. Science 323:373-375.

Moehring, A. J., A. Llopart, S. Elwyn, J. A. Coyne, and T. F. Mackay. 2006. The genetic basis of prezygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to mating preference. Genetics 173:215-223.

Moehring, A. J., J. Li, M. D. Schug, S. G. Smith, M. deAngelis, T. F. Mackay, and J. A. Coyne. 2004. Quantitative trait loci for sexual isolation between *Drosophila simulans* and *D. mauritiana*. Genetics 167:1265-1274.

Moehring, A. J., and T. F. Mackay. 2004. The quantitative genetic basis of male mating behavior in *Drosophila melanogaster*. Genetics 167:1249-1263.

Nanda, P., and B. N. Singh. 2011. Effect of chromosome arrangements on mate recognition system leading to behavioral isolation in *Drosophila ananassae*. Genetica 139:273-279.

Noor, M. A., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. Prac. Natl. Acad. Sci. USA 98:12084-12088.

O'Dell, K. M., J. D. Armstrong, M. Y. Yang, and K. Kaiser. 1995. Functional dissection of the *Drosophila* mushroom bodies by selective feminization of genetically defined subcompartments. Neuron 15:55-61.

Ortiz-Barrientos, D., B. A. Counterman, and M. A. Noor. 2004. The genetics of speciation by reinforcemnt. PLoS Biol. 2:e416.

Robertson, H. M. 1982. Female courtship summation in *Drosophila melanogaster*. Anim. Behav. 30:1105-1117.

Roll-Mecak, A., and F. J. McNally. 2010. Microtuble-severing enzymes. Curr. Opin. Cell Biol. 22:96-103.

Ruedi, E. A., and K. A. Hughes. 2009. Natural genetic variation in complex mating behaviors of male *Drosophila melanogaster*. Behav. Genet. 38:424-436.

Sawamura, K., H. Zhi, K. Setoguchi, H. Yamada, T. Miyo, M. Matsuda, and Y. Oguma. 2008. Genetic analysis of female mating recognition between *Drosophila ananassae* and *Drosophila pallidosa*: application of interspecific mosaic genome lines. Genetica 133:179-185.

Souse-Neves, R., and A. Rosas. 2010. An analysis of genetic changes during the divergence of *Drosophila* species. PLoS ONE 5:e10485.

Suzuki, K., N. Juni, and D. Yamamoto. 1997. Enhanced mate refusal in female *Drosophila* induced by a mutation in the *spinster* locus. Appl. Entomol. Zool. 32:235-243.

Uenoyama, T., and Y. Inoue. 1995. Genetic studies on premating isolation in *Drosophila simulans*. I.A *D. simulans* line highly crossable with *D. melanogaster*. Jpn. J. Genet. 70:365-371.

Wheeler, D. A., C. P. Kyriacou, M. L. Greenacre, Q. Yu, J. E. Rutila, M. Rosbash, and J. C. Hall. 1991. Molecular transfer of a species-specific behaviour from *Drosophila simulans* to *Drosophila melanogaster*. Science 251:1082-1085.

Yamamoto D., and Y. Nakano. 1999. Sexual behavior mutants revisited: molecular and cellular basis of *Drosophila* mating. Cell. Mol. Life Sci. 56:634-646.

Zhang, D., K. D. Grode, S. F. Stewman, J. D. Diaz-Valencia, E. Liebling, T. Rath, T. Riera, J. D. Currie, D. W. Buster, A. B. Asenjo, H. J. Sosa, J. L. Ross, A. Ma, S. L. Rogers, and D. J. Sharp. 2011. *Drosophila katanin* is a microtubule depolymerase that regulated cortical-microtubule plus-end interactions and cell migration. Nat. Cell. Biol. 13:361-370.

Zhang, D. S., and W. F. Odenwald. 1995. Misexpression of the *white* (*w*) gene triggers male-male courtship in *Drosophila*. Proc. Natl. Acad. Sci. USA 92:5525-5529.

Chapter 4: General Discussion

This study identified five regions, and fine mapped four, on the right arm of the third chromosome and identified the first candidate gene for female preference, interspecific female preference, and behavioural isolation in *Drosophila*. This was accomplished through deficiency mapping and took advantage of the genetic tools available in *D. melanogaster* such as a sequenced genome, readily available genetic mutant lines, and a database outlining each gene and its known function (FlyBase online database; www.flybase.org; Tweedie *et al.* 2009). With use of this species, additional fine mapping and the identification the candidate gene in each region can be complete with relative ease.

This study mapped genes for interspecific female preference between *D. simulans* and *D. melanogaster* on the right arm of the third chromosome because this region (or the equivalent of it in other species) has been repeatedly found to influence behavioural isolation. When mapping the whole genome, previous studies have identified the entire third chromosome (Coyne 1992; Coyne 1989), the right arm (Ting *et al.* 2001; Noor *et al.* 2001), or specific regions with in 3R (Doi *et al.* 2001; Moehring *et al.* 2004; Moehring *et al.* 2006) as regions that most likely hold genes for interspecific female preference. Due to the use of *D. melanogaster* and the type of mapping that I used, the regions identified in this study were relatively small compared to the those previously identified. The regions identified here are spread along the entire arm which may explain why whole arms were easily identified in other research. However, some studies identified only a single region on the

chromosome (Doi *et al.* 2001; Noor *et al.* 2001; Moehring *et al.* 2006) and not multiple regions scattered throughout as I have found. However, these studies may have simply identified that regions with the most effect and were not able to identify gene(s) in other regions with only minor influence on the trait.

Of the two studies that used QTL mapping to investigate interspecific female preference, some but not all of the regions identified overlapped the regions identified in this study. Moehring *et al.* (2006), which investigated the genetic basis of *D. santomea* females rejecting *D. yakuba* males identified one regions at 82A-88B with a peak at 85E. Although the peak falls within a region that I was not able to map, the region did overlap with the region found in this study which contains the candidate gene *katanin-60*. Moehring *et al.* (2004), which investigated the genetic basis of *D. mauritiana* females rejecting *D. simulans* males identified three regions. One region and its peak, 88B-93F with a peak at 91C, overlapped a region found in this study, and another region overlapped but its peak did not (97D-100E with a peak at 99F). Although both regions identified 88B, I did not find this region to hold genes for interspecific female preference in this species pair.

It is possible that some of the genes important for interspecific female preference act to isolate many different species pairs. However, not all regions identified in this study and in a previous study (Moehring *et al.* 2004) have been found in other mapping studies. This may be due to the low number of attempts to map female preference genes to specific loci or the reduced genetic variability within inbred lines used in most genetic research. In theory, mapping the genetic basis of this trait in each line of each species may reveal that females from different

species are isolated from heterospecific males due to different variants of the same gene. However, it is more likely that the genes for interspecific female preference]differ for each species or even for each species pair but may be involved in the same biological processes or pathways.

Although I identified a candidate gene, *katanin-60*, future work is required to confirm this as a gene for behavioural isolation for this species pair. Further investigations should also include finer mapping of the five regions identified in this study, the identification of the other variants responsible for this trait, and an examination into the role of *katanin-60* in species isolation in other species pairs of *Drosophila*.

With addition to 3R, I also tested two genes previously identified to influence *D. melanogaster* intraspecific female preference (*spinster*, Suzuki *et al.* 1997; *dissatisfaction*, Finley *et al.* 1997). However, neither of these genes were found to have an effect on interspecific female preference in *D. simulans*, and many genes on 3R shown to influence male intraspecific mating behaviour were also not within our regions, such as *fruitless* (Hall 1978) and *voila* (Balakireva *et al.* 1998). One explanation is that the genes that influence intraspecific mating behaviour in one species do not influence mating in others: the genes found through mutation studies were identified in *D. melanogaster* and I mapped genes in *D. simulans*. The gene *period* was originally identified to influence male courtship song in *D. melanogaster* and has also been found to control *D. simulans* male courtship song as well (Wheeler *et al.* 1991). Therefore, it is likely that genes found to influence *D. melanogaster* mating behaviour do also influence *D. simulans* intraspecific mating. Another more

likely alternative is that the genetic basis of intra- and interspecific mating behaviour are not one and the same. Other studies that have identified regions within the genome that contributed to interspecific female preference in other *Drosophila* species pairs also did not find genes found to influence intraspecific mating (Moehring *et al.* 2004; Moehring *et al.* 2006). These genes identified through mutagenesis may not contribute to this trait because they do not vary in the population (Ruedi and Hughes 2009) and do not contribute to variation of mating behaviour within the species (Moehring and Mackay 2004; Mackay et al. 2005).

The pursuit of identifying the genetic basis of behavioural isolation in *Drosophila* has provided insight into what is happening on the genetic level in order to maintain species isolation. However, understanding this process can also aid in the understanding of the process of speciation as this thought to be the first mechanism to evolve to stop gene flow between two population (Coyne and Orr 1997). As the genes that stop these two species, *D. melanogaster* and *D. simulans*, are identified and confirmed for 3R and the rest of the genome, together with the genetic basis for postzygotic isolation and post-copulatory prezygotic isolation will paint a complete picture of how species are formed and maintained. We are at the cusp of identifying individual genes that influence behavioral isolation, and thus an understanding of the primary genetic basis of speciation.

References

Barnwell, C. V., and M. A. Noor. 2008. Failure to replicate two mate preference QTLs across multiple strains of *Drosophila pseudoobscura*. J. Hered. 99:653-656.

Carracedo, M. C., A. Suarez, A. Asenjo, and P. Casares. 1998. Genetics of hybridization between *Drosophila simulans* females and *D. melanogaster* males. Heredity 80:17-24.

Coyne, J. A. 1989. Genetics of sexual isolation between two sibling species, *Drosophila simulans* and *Drosophila mauritiana*. Proc. Natl. Acad. Sci USA 86:5464-5468.

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

Coyne, J. A., and H. A. Orr. 1997. Patterns of speciation in *Drosophila* revisited. Evolution 51: 295-303.

Doi, M., M. Matsuda, M. Tomaru, H. Matsubayashi, and Y. Oguma. 2001. A locus for female discrimination behavior causing sexual isolation in *Drosophila*. Proc. Natl. Acad. Sci. USA 98:6714-6719.

Finley, K. D., B. J. Taylor, M. Milstein, and M. McKeown. 1997. *dissatisfaction*, a gene involved in sex-specific behavior and neural development of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 94:913-918.

Mackay, T. F., S. L. Heinsohn, R. F. Lynman, A. J. Moehring, T. J. Morgan, and S. M. Rollmann. 2005. Genetics and Genomics of *Drosophila* mating behavior. Prac. Natl. Acad. Sci. USA 102:6622-6629.

Moehring, A. J., A. Llopart, S. Elwyn, J. A. Coyne, and T. F. Mackay. 2006. The genetic basis of prezygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to mating preference. Genetics 173:215-223.

Moehring, A. J., and T. F. Mackay. 2004. The quantitative genetic basis of male mating behavior in *Drosophila melanogaster*. Genetics 167:1249-1263.

Noor, M. A., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. Prac. Natl. Acad. Sci. USA 98:12084-12088.

Ortiz-Barrientos, D., B. A. Counterman, and M. A. Noor. 2004. The genetics of speciation by reinforcemnt. PLoS Biol. 2:e416.

Ruedi, E. A., and K. A. Hughes. 2009. Age, but not experience, affects courtship gene expression in male *Drosophila melanogaster*. PLoS one.

Sawamura, K., H. Zhi, K. Setoguchi, H. Yamada, T. Miyo, M. Matsuda, and Y. Oguma. 2008. Genetic analysis of female mating recognition between *Drosophila ananassae* and *Drosophila pallidosa*: application of interspecific mosaic genome lines. Genetica 133:179-185.

Suzuki, K., N. Juni, and D. Yamamoto. 1997. Enhanced mate refusal in female *Drosophila* induced by a mutation in the *spinster* locus. Appl. Entomol. Zool. 32:235-243.

Uenoyama, T., and Y. Inoue. 1995. Genetic studies on premating isolation in *Drosophila simulans*. I.A *D. simulans* line highly crossable with *D. melanogaster*. Jpn. J. Genet. 70:365-371.

Wheeler, D. A., C. P. Kyriacou, M. L. Greenacre, Q. Yu, J. E. Rutila, M. Rosbash, and J. C. Hall. 1991. Molecular transfer of a species-specific behaviour from *Drosophila simulans* to *Drosophila melanogaster*. Science 251:1082-1085.

Appendix A: List of the candidate genes in the five regions on 3R

Candidate Region 82E7-F7

Gene Name	Region Cytological	Sequence	Gene Molecular Function	Gene Biological Process
CR43635	82E6	870927..871670	unknown	unknown
corto	82E7	904244..912749	protein binding; protein homodimerization activity	chromatin silencing
CG43131	82F1	934400..934900	unknown	unknown
CG12007	82F1	953810..955665	Rab geranylgeranyltransferase activity (predicted)	regulation of cell shape; cell adhesion; neurogenesis
CG12589	82F1	962779..963295	unknown	unknown
dpr16	82F1-3	976630..995849	unknown	unknown
CG12590	82F1	985334..986627	unknown	unknown
cno	82F4-6	998,392..1,043,147	actin binding (predicted)	embryonic morphogenesis; dorsal closure; regulation of JNK cascade epidermis morphogenesis; compound eye development
snRNA:U4atac:82E	82F4	1,020,726..1,020,885	unknown	nuclear mRNA splicing, via spliceosome; RNA splicing (predicted)
Probeta2R2	82F6	1,043,473..1,045,168	endopeptidase activity (predicted)	proteolysis involved in cellular protein catabolic process (predicted)
CG1116	82F6	1,045,390..1,047,270	unknown	unknown
CG2604	82F6	1,047,354..1,049,197	nucleotide binding; oxidoreductase activity (predicted)	bristle development; wing disc development
CG14664	82F6	1,049,861..1,050,492	unknown	unknown
CG1115	82F6	1,051,882..1,053,248	unknown	unknown

katanin-60	82F6	1,053,011..1,057,940	protein binding	bristle development;wing disc development
Mms19	82F6	1,058,268..1,062,011	binding	DNA repair
rtp	82F6	1,061,814..1,063,037	unknown	engulfment of apoptotic cell
CG12163	82F6-7	1,063,373..1,077,468	cysteine-type endopeptidase activity (predicted)	bristle development; wing disc development; autophagic and salivary gland cell death
CG1113	82F7	1,079,070..1,081,125	unknown	unknown
CG12173	82F7	1,081,166..1,082,423	acireductone synthase activity; 2,3-diketo-5-methylthiopentyl-1-phosphate enolase activity; magnesium ion binding; phospho-glycolate phosphatase activity; 2-hydroxy-3-keto-5-methylthio-pentenyl-1-phosphate activity (predicted)	L-methionine salvage from methylthioadenosine (predicted)
Hph	82F7-8	1,082,763..1,094,197	peptidyl-proline 4-dioxygenase activity (predicted)	response to hypoxia; protein localization; response to DNA damage stimulus
Tim17a2	82F7-8	1,084,058..1,088,414	P-P-bond-hydrolysis-driven protein transmembrane transporter activity (predicted)	protein targeting to mitochondrion; protein transport (predicted)

Candidate Region 89B10-16

Gene Name	Region Cytological	Sequence	Gene Molecular Function	Gene Biological Process
gish	89B9-12	12,098,176..12,128,443	protein serine/threonine kinase activity (predicted)	olfactory learning; glial cell migration; protein phosphorylation; Wnt receptor signaling pathway; spermatogenesis
Zip3	89B12	12,131,485..12,134,424	metal ion transmembrane transporter activity	transmembrane transport; metal ion transport (predicted)
Sulf1	89B12-13	12,136,264..12.164,260	N-acetylglucosamine-6- sulfatase activity; alkyl sulfatase activity (predicted)	regulation of Wnt receptor signaling pathway; pattern specification process regulation of transforming growth factor beta receptor signaling pathway
tRNA:F2:89BC	89B12	12,147,125..12,147,197	UUC codon-amino acid adaptor activity (predicted)	translation
tRNA:V4:89BC	89B12	12,147,412..12,147,484	GUU codon-amino acid adaptor activity (predicted)	translation
CG6901	89B12	12,149,137..12,150,850	transmembrane transporter activity (predicted)	transmembrane transporter (predicted)
CG17929	89B12	12,151,214..12,153,053	transmembrane transporter activity (predicted)	transmembrane transport (predicted)
CG17930	89B12	12,153,572..12,155,344	transmembrane transporter activity (predicted)	transmembrane transport (predicted)
SF2	89B13	12,164,665..12,167,706	protein binding	mitotic cell cycle G2/M transition DNA damage checkpoint; Neurogenesis

pad	89B13	12,168,040..12,171,551	zinc ion binding; nucleic acid binding (predicted)	bristle morphogenesis
CG14879	89B13	12,171,855..12,174,946	sugar binding (predicted)	unknown
Manf	89B13	12,175,088..12,176,523	unknown	neuron homeostasis; neuron projection development
CG10311	89B13	12,176,627..12,182,733	unknown	unknown
CG17931	89B13	12,183,866..12,185,830	unknown	unknown
CG42446	89B13	12,183,866..12,185,038	unknown	unknown
CG14880	89B13	12,186,051..12,194,026	chitin binding (predicted)	chitin metabolic process (predicted)
CG10317	89B13-14	12,195,093..12,196,255	unknown	unknown
ss	89B14-15	12,200,148..12,229,412	sequence-specific DNA binding	regulation of dendrite morphogenesis;antennal development; antennal morphogenesis; leg segmentation
tRNA:CR31497	89B16	12,250,709..12,250,782	ACU codon-amino acid adaptor binding (predicted)	translation
CG31279	89B16	12,256,527..12,258,619	unknown	unknown
tRNA:CR31282	89B16	12257662..12257735	ACU codon-amino acid adaptor activity (predicted)	translation
CG17565	89B16	12,258,703..12,260,462	Rab geranylgeranyl-transferase activity (predicted)	unknown
CG14881	89B16	12,260,647..12,261,788	binding; oxidoreductase activity (predicted)	metabolic process (predicted)
mRpS33	89B16	12261767..12262315	structural constituent of ribosome (predicted)	translation (predicted)
ema	89B16	12262602..12267382	unknown	endosome to lysosome transport; endosome transport
CG14894	89B16	12,267,728..12,268,729	binding (predicted)	unknown
CG14882	89B16	12,268,966..12,270,526	reductase activity (predicted)	oxidation-reduction process

Candidate Region 89D6-E2

Gene Name	Region Cytological	Sequence	Gene Molecular Function	Gene Biological Process
nonA-l	89D6	12,465,929..12,468,541	poly-pyrimidine tract binding; mRNA binding (predicted)	nuclear mRNA splicing, via spliceosome (predicted)
CG10326	89D6	12,469,327..12,470,887	unknown	unknown
CG10324	89D6	12,470,598..12,472,711	nucleic acid binding (predicted)	neurogenesis
Cctgamma	89D6	12,473,311..12,475,700	ATPase activity, coupled; unfolded protein binding (predicted)	mitotic spindle organization
CG14906	89D6	12,475,703..12,478,029	methyltransferase activity; nucleic acid binding (predicted)	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process; methylation (predicted)
CG14907	89D6	12,475,703..12,478,029	unknown	unknown
modSP	89D6	12,478,168..12,481,489	serine-type endopeptidase activity	innate immune response
Ubx	89D6	12,482,345..12,560,348	protein binding; sequence-specific DNA binding; protein domain specific binding; DNA bindin	anatomical structure development; biological regulation; regulation of multicellular organismal development; regulation of biological process; organ development; regulation of developmental process; cell fate commitment; cell fate determination; segment specification; organ morphogenesis
bxd	89E1	12,567,847..12,598,911	unknown	unknown
CG31275	89E1	12,576,454..12,577,111	ATPase activity, coupled (predicted)	lateral inhibition

tre-3	89E1	12,589,091..12,589,441	SET domain binding	establishment of protein localization to chromatin; positive regulation of transcription, DNA dependent
tre-2	89E1	12,589,406..12,590,514	SET domain binding	establishment of protein localization to chromatin; positive regulation of transcription, DNA dependent
tre-1	89E1	12,591,129..12,592,078	SET domain binding	establishment of protein localization to chromatin; positive regulation of transcription, DNA dependent
Glut3	89E1	12,614,033..12,615,745	glucose transmembrane transporter activity (predicted)	glucose transport
abd-A	89E2	12,632,936..12,655,767	sequence-specific enhancer binding RNA polymerase II transcription factor activity; sequence-specific DNA binding; sequence-specific DNA binding transcription factor activity	anatomical structure development; cellular process; multicellular organismal development; biological regulation; regulation of cellular macromolecule biosynthetic process; anterior/posterior pattern formation, imaginal disc; cardioblast differentiation; embryonic morphogenesis; cellular process involved in reproduction; gonadal mesoderm development; neuroblast fate commitment

Candidate Region 91B2-C4

Gene Name	Region Cytological	Sequence	Gene Molecular Function	Gene Biological Process
CG7691	91B4	14,385,591..14,387,075	zinc ion binding; nucleic acid binding (predicted)	unknown
CG43210	91B4	14,390,214..14,390,895	unknown	unknown
CG43194	91B4	14,390,214..14,390,895	unknown	unknown
CG7694	91B4-5	14,398,817..14,416,448	zinc ion binding (predicted)	unknown
fray	91B4-5	14,398,817..14,415,175	protein serine/threonine kinase activity (predicted)	establishment of blood-nerve barrier; axonal fasciculation; axon ensheathment; chitin-based embryonic cuticle biosynthetic process
CG14306	91B5	14,416,822..14,418,600	zinc ion binding (predicted)	unknown
CG14304	91B5-6	14,418,523..14,447,049	chitin binding (predicted)	chitin metabolic process (predicted)
CG14305	91B5	14,424,478..14,425,819	protein kinase activity	protein phosphorylation
CG7695	91B6	14,442,476..14,443,233	unknown	unknown
tRNA:CR31228	91B7	14,454,189..14,454,260	CCA codon-amino acid adaptor activity (predicted)	translation
CG14301	91B7	14,454,870..14,467,179	chitin binding (predicted)	chitin metabolic process (predicted)
CG14303	91B7-8	14,470,177..14,475,859	unknown	unknown
CG31224	91B8	14,475,843..14,484,963	nucleic acid binding (predicted)	unknown
CstF-64	91B8	14,485,114..14,486,850	mRNA binding (predicted)	neurogenesis
Cpsf73	91B8	14,486,843..14,489,250	DNA binding	mRNA polyadenylation; histone mRNA 3'-end processing; neurogenesis

Gos28	91B8	14,489,464..14,491,029	SNAP receptor activity (predicted)	ER to Golgi vesicle-mediated transport; intra-Golgi vesicle mediated transport; vesicle mediated transport (predicted)
CG31231	91B8	14,491,083..14,493,455	unknown	unknown
CG31229	91B8	14,493,595..14,494,805	P-P-bond-hydrolysis-driven protein transmembrane transporter activity (predicted)	protein targeting to mitochondrion (predicted)
CG7702	91B8-C1	14,495,114..14,500,401	unknown	unknown
CG31230	91B8	14,497,401..14,497,816	unknown	unknown
CG7705	91C1	14,500,376..14,503,278	unknown	unknown
CG7706	91C1	14,503,605..14,506,043	anion exchanger adaptor activity (predicted)	unknown
CG7708	91C1	14,506,328..14,511,029	amino acid transmembrane transporter activity; proline:sodium symporter activity; choline transmembrane transporter activity (predicted)	acetylcholine biosynthetic process; amino acid transmembrane transport (predicted)
Muc91C	91C1	14,513,308..14,518,441	unknown	neurogenesis
CG14300	91C1	14,519,015..14,519,428	unknown	unknown
CG34282	91C1	14,520,974..14,521,498	unknown	unknown

Candidate Region 93B7-C5

Gene Name	Region Cytological	Sequence	Gene Molecular Function	Gene Biological Process
Dhc93AB	93B5-7	16,850,741..16,868,141	ATPase activity, coupled; motor activity (predicted)	microtubule-based movement (predicted)
CG12278	93B7	16,869,815..16,870,627	unknown	unknown
CG31189	93B7	16,870,620..16,871,533	unknown	unknown
CG31207	93B7	16,871,754..16,872,749	unknown	unknown
CG7079	93B7	16,873,487..16,874,543	unknown	unknown
CG17279	93B7	16,875,736..16,876,637	unknown	unknown
Mvl	93B7-8	16,877,105..16,886,517	manganese ion transmembrane transporter activity; iron ion transmembrane transporter activity; copper ion transmembrane transporter activity; symporter activity	iron assimilation; multicellular organismal iron ion homeostasis; copper ion transport; divalent metal ion transport; copper ion import; sensory perception of sweet taste; entry of virus into host cell; transition metal ion transport; copper ion homeostasis; transition metal ion homeostasis
Cortactin	93B8	16,887,402..16,890,193	proline-rich region binding	regulation of cell shape; positive regulation of receptor-mediated endocytosis; border follicle cell migration; female germline ring canal formation
AnnIX	93B9-10	16,890,951..16,896,639	calcium ion binding; calcium-dependent phospholipid binding; actin binding (predicted)	wing disc dorsal/ventral pattern formation

r-l	93B10	16,896,275..16,899,039	orotidine-5'-phosphate decarboxylase activity; orotate phosphoribosyl-transferase activity	de novo' pyrimidine base biosynthetic process
dmrt93B	93B10	16,899,729..16,902,520	sequence-specific DNA binding transcription factor activity (predicted)	sex differentiation; regulation of transcription, DNA-dependent (predicted)
CG7056	93B10	16,902,635..16,906,405	sequence-specific DNA binding transcription factor activity (predicted)	dendrite morphogenesis; neurogenesis
RhoGAP93B	93B10-11	16,908,758..16,923,527	unknown	axon guidance
CG7044	93B11-12	16,923,892..16,927,436	binding (predicted)	unknown
CG5745	93B12	16,927,936..16,930,835	GTPase activator activity (predicted)	phagocytosis, engulfment
sec15	93B12	16,930,943..16,933,686	unknown	bristle development; phototaxis; border follicle cell migration; axon guidance; endocytic recycling
rtet	93B12	16,934,170..16,936,554	sugar transmembrane transporter activity (predicted)	oogenesis
Rab11	93B12-13	16,937,129..16,941,877	GTPase activity (predicted)	cellular component organization or biogenesis; sensory organ development; multicellular organism reproduction; biological regulation;localization; gamete generation; organelle organization; cell cycle; regulation of developmental process; cell cycle process; fusome organization

ppan	93B13	16,941,958..16,943,621	unknown	imaginal disc development; larval development; oogenesis; neurogenesis
CG17282	93B13	16,943,579..16,944,571	binding (predicted)	unknown
slmb	93B13-C1	16,944,973..16,953,202	phosphoprotein binding	biological regulation; multicellular,organism reproduction; ovarian follicle cell development; gamete generation; cellular component organization or biogenesis; localization; learning or memory; rhythmic process; regulation of cellular component organization; system process
CG5793	93C1	16,953,295..16,955,323	catalytic activity (predicted)	metabolic process (predcited)
Obp93a	93C1	16,954,672..16,955,439	odorant binding (predicted)	sensory perception of chemical stimulus (predicted)
CG10825	93C1	16,955,941..16,958,442	unknown	unknown
CG7009	93C1	16,958,456..16,959,599	rRNA (uridine-2'-O-)-methyltransferase activity (predcited)	rRNA modification (predicted)
Ubpy	93C1	16,959,984..16,963,936	ubiquitin thiolesterase activity (predicted)	positive regulation of canonical Wnt receptor signaling pathway; imaginal disc-derived wing margin morphogenesis
CG5802	93C1	16,963,992..16,966,087	UDP-galactose transmembrane transporter activity (predicted)	transmembrane transport (predicted)
SNF4Agamma	93C1-5	16,966,463..17,038,409	AMP-activated protein kinase activity (predicted)	positive regulation of cell cycle; cholesterol homeostasis
CG10824	93C2	16,984,960..16,986,815	unknown	unknown
CG5810	93C2	16,989,512..16,990,798	unknown	unknown

Snmp1	93C2	16,991,220..16,993,927	scavenger receptor activity (predicted)	response to pheromone; cell surface receptor linked signaling pathway
CG5862	93C6	17,040,782..17,041,948	unknown	unknown
CG3353	93C6	17,042,988..17,044,511	zinc ion binding (predicted)	unknown

Candidate Region 99F2-100C6

Gene Name	Region Cytological	Sequence	Gene Molecular Function	Gene Biological Process
Fer1HCH	99F2	26,211,295..26,213,851	ferrous iron binding	cellular iron ion homeostasis
Fer2LCH	99F2	26,213,554..26,216,306	ferrous iron binding	bristle development; cellular iron ion homeostasis
CG2217	99F3-4	26,244,489..26,248,693	unknown	unknown
CG42740	99F3-4	26,244,489..26,248,693	unknown	unknown
CG15535	99F4	26,248,754..26,249,565	unknown	unknown
mRpS18C	99F4	26,249,696..26,250,278	structural constituent of ribosome (predicted)	translation (predicted)
CG15536	99F4	26,250,221..26,250,884	unknown	unknown
CG2218	99F4	26,251,382..26,253,998	ubiquitin-protein ligase activity (predicted)	protein ubiquitination (predicted)
CG15533	99F4	26,254,548..26,257,091	sphingomyelin phosphodiesterase activity (predicted)	sphingomyelin metabolic process (predicted)

CG15534	99F4	26,257,786..26,260,861	sphingomyelin phosphodiesterase activity (predicted)	sphingomyelin metabolic process (predicted)
CG15537	99F4	26,261,033..26,269,321	hormone binding (predicted)	hormone metabolic process (predicted)
Osi23	99F4	26,269,871..26,271,031	unknown	unknown
aralar1	99F4-5	26,271,357..26,281,589	calcium ion binding; transmembrane transporter activity (predicted)	mitochondrial transport (predicted)
CG2224	99F5	26,282,683..26,285,110	unknown	unknown
CDase	99F5-6	26,285,511..26,291,458	ceramidase activity	synaptic transmission; sphingolipid metabolic process; photoreceptor cell maintenance; ceramide catabolic process; synaptic vesicle fusion to presynaptic membrane; hatching behavior; synaptic vesicle exocytosis
PH4aEFB	99F6	26,291,774..26,311,521	procollagen-proline 4-dioxygenase activity (predicted)	peptidyl-proline hydroxylation to 4- hydroxy-L-proline (predicted)
spdo	99F6	26,302,411..26,304,571	Notch binding; protein binding	negative regulation of Notch signaling pathway; heart morphogenesis; sensory organ development; Notch signaling pathway; sensory organ precursor cell division; asymmetric cell division
Jon99Fii	99F6	26,312,735..26,313,648	endopeptidase activity	proteolysis
Jon99Fi	99F6	26,314,687..26,315,625	serine-type endopeptidase activity (predicted)	proteolysis (predicted)

PH4aSG2	99F6	26,315,798..26,317,724	procollagen-proline 4-dioxygenase activity (predicted)	salivary gland morphogenesis
PH4aMP	99F6	26,317,535..26,321,875	procollagen-proline 4-dioxygenase activity (predicted)	peptidyl-proline hydroxylation to 4-hydroxy-L-proline (predicted)
PH4aNE1	99F6	26,322,473..26,326,048	procollagen-proline 4-dioxygenase activity (predicted)	peptidyl-proline hydroxylation to 4-hydroxy-L-proline (predicted)
CG31371	99F6	26,326,924..26,329,082	binding; oxidoreductase activity acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen; procollagen-proline 4-dioxygenase activity (predicted)	oxidation-reduction process (predicted)
PH4aG1	99F6	26,329,387..26,331,371	procollagen-proline 4-dioxygenase activity (predicted)	salivary gland morphogenesis
CG15539	99F6	26,331,709..26,333,699	procollagen-proline 4-dioxygenase activity (predicted)	peptidyl-proline hydroxylation to 4-hydroxy-L-proline (predicted)
mir-4908	99F6	26,332,551..26,332,609	unknown	unknown
CG34041	99F6	26,333,745..26,335,868	unknown	unknown
PH4aNE2	99F6	26,335,805..26,337,936	procollagen-proline 4-dioxygenase activity (predicted)	peptidyl-proline hydroxylation to 4-hydroxy-L-proline (predicted)
CG31524	99F6-7	26,338,044..26,340,077	procollagen-proline 4-dioxygenase activity (predicted)	oxidation-reduction process (predicted)
CG9698	99F7	26,340,546..26,342,641	procollagen-proline 4-dioxygenase activity (predicted)	peptidyl-proline hydroxylation to 4-hydroxy-L-proline (predicted)
jdp	99F7	26,342,774..26,352,761	heat shock protein binding; unfolded protein binding (predicted)	protein folding (predicted)
tmod	99F7-8	26,358,336..26,404,627	actin binding (predicted)	cytoskeleton organization (predicted)

CG34155	99F8	26,380,276..26,385,921	unknown	unknown
CG9702	99F9	26,405,151..26,407,839	high affinity sulfate transmembrane transporter activity (predicted)	transmembrane transport (predicted)
Rpt6R	99F9-10	26,407,920..26,409,419	ATPase activity (predicted)	neurogenesis
CG9717	99F10	26,409,606..26,413,021	high affinity sulfate transmembrane transporter activity (predicted)	transmembrane transport (predicted)
CG2246	99F10-100A1	26,418,691..26,427,211	ribose phosphate diphosphokinase activity (predicted)	nucleotide biosynthetic process (predicted)
CG31019	99F11-100A1	26,422,726..26,425,018	zinc ion binding; metallocarboxy-peptidase activity (predicted)	proteolysis (predicted)
CG31021	100A1	26,427,928..26,429,763	binding; oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen; procollagen-proline 4-dioxygenase activity (predicted)	oxidation-reduction process (predicted)
PH4aNE3	100A1	26,429,903..26,431,866	procollagen-proline 4-dioxygenase activity (predicted)	oxidation-reduction process (predicted)
CG31016	100A1	26,431,911..26,436,178	procollagen-proline 4-dioxygenase activity (predicted)	peptidyl-proline hydroxylation to 4-hydroxy-L-proline (predicted)
CG2267	100A1	26,452,962..26,454,750	unknown	unknown
CG31013	100A1	26,469,489..26,471,602	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen; iron ion binding; L-ascorbic acid binding; procollagen-proline 4-dioxygenase activity (predicted)	oxidation-reduction process (predicted)

PH4alphaPV	100A1	26,471,828..26,474,273	procollagen-proline 4-dioxygenase activity (predicted)	peptidyl-proline hydroxylation to 4-hydroxy-L-proline (predicted)
CG34432	100A1	26,474,873..26,475,595	unknown	unknown
CG34433	100A1	26,475,683..26,476,580	unknown	unknown
Spn100A	100A2	26,516,366..26,518,490	serine-type endopeptidase inhibitor activity (predicted)	unknown
CG12069	100A2	26,519,294..26,521,220	protein serine/threonine activity	protein phosphorylation kinase
Pka-C2	100A2	26,521,258..26,522,860	cAMP-dependent protein kinase activity (predicted)	neurogenesis
CG31010	100A2	26,530,481..26,531,815	unknown	unknown
CG1340	100A3	26,551,809..26,553,719	mRNA binding (predicted)	translational initiation (predicted)
CG11313	100A3	26,555,559..26,567,010	serine-type endopeptidase activity (predicted)	proteolysis
CG15543	100A3	26,568,602..26,569,637	nucleobase, nucleoside, nucleotide kinase activity; ATP binding (predicted)	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process (predicted)
Npc2g	100A3	26,570,326..26,570,983	sterol binding (predicted)	mesoderm development
Npc2h	100A3	26,571,201..26,572,010	sterol binding (predicted)	sterol transport (predicted)
CR43238	100A3	26,572,407..26,572,894	unknown	unknown
zfh1	100A4-5	26,591,648..26,614,205	DNA binding	motor axon guidance; hemocyte development; nervous system development; antimicrobial humoral response; germ cell migration; garland cell differentiation; lymph gland development; mesoderm development

wts	100A5	26,615,379..26,632,341	protein binding	biological regulation; anatomical structure development; sensory organ development; organ development; regulation of developmental process; compound eye development; cell proliferation; response to ionizing radiation; growth; cell cycle
mir-1013	100A5	26,617,358..26,617,418	unknown	unknown
dj-1beta	100A5	26,632,848..26,633,701	unknown	response to oxidative stress; adult locomotory behavior
cindr	100A6	26,634,349..26,652,542	SH3 domain binding	positive regulation of receptor mediated endocytosis; border follicle cell migration; compound eye development; lateral inhibition; intercellular bridge organization; cytokinesis; actin filament organization; compound eye morphogenesis
CG15544	100A6	26,653,572..26,667,756	unknown	unknown
tll	100A6	26,678,037..26,680,095	DNA binding	torso signaling pathway; terminal region determination; ring gland development; optic lobe placode development; regulation of cell cycle; cell fate commitment; neuroblast division
Cpr100A	100A6	26,693,245..26,694,908	structural constituent of chitin-based cuticle (predicted)	unknown
CG15545	100A6	26,696,780..26,698,002	unknown	unknown
CG15546	100A6	26,700,087..26,701,908	unknown	unknown

CG12071	100A6	26,702,806..26,706,597	zinc ion binding; nucleic acid binding (predicted)	phagocytosis, engulfment
CG15547	100A6	26,707,514..26,709,225	nucleoside diphosphate kinase activity; ATP binding (predicted)	nucleoside diphosphate phosphorylation; GTP biosynthetic process; CTP biosynthetic process;UTP biosynthetic process (predicted)
Sap-r	100A6-7	26,709,186..26,714,795	unknown	dsRNA transport
Cyp4c3	100A7	26,719,683..26,726,410	electron carrier activity	oxidation-reduction process (predicted)
CG33483	100A7	26,726,440..26,727,883	unknown	unknown
Ptx1	100A7-B1	26,738,610..26,756,404	sequence-specific DNA binding transcription factor activity; sequence-specific DNA binding (predicted)	dendrite morphogenesis
CG15549	100B1	26,761,118..26,762,229	growth factor activity (predicted)	neurogenesis
CG15550	100B1	26,762,252..26,762,824	unknown	unknown
CG15548	100B1	26,785,529..26,787,973	unknown	unknown
Ctr1C	100B1	26,789,011..26,790,042	copper ion transmembrane transporter activity (predicted)	copper ion transmembrane transport (predicted)
5-HT7	100B1	26,794,864..26,842,569	G-protein coupled amine receptor activity; serotonin receptor activity (predicted)	male courtship behavior, proboscis-mediated licking; female mating behavior; male mating behavior; courtship behavior; male courtship behavior, orientation prior to leg tapping and wing vibration; male courtship behavior, veined wing vibration
CG31008	100B1	26,815,199..26,815,728	unknown	unknown

CG31007	100B1	26,817,457..26,818,179	unknown	lateral inhibition
CG33773	100B1	26,861,068..26,861,846	unknown	unknown
CanA1	100B1	26,865,098..26,872,543	calcium-dependent protein serine/ threonine phosphatase activity; protein serine/threonine phosphatase activity (predicted)	protein dephosphorylation (predicted)
CG33920	100B1	26,870,113..26,870,827	unknown	unknown
Aph-4	100B1	26,874,654..26,878,268	alkaline phosphatase activity (predicted)	epithelial fluid transport; nervous system development
mir-4948	100B1	26,876,820..26,876,925	unknown	unknown
l(3)03670	100B1	26,878,333..26,879,634	unknown	unknown
dco	100B2	26,880,905..26,886,931	kinase activity	regulation of biological process; anatomical structure development; phosphorus metabolic process; macromolecule modification; growth; cellular component organization or biogenesis; response to cocaine; rhythmic process; establishment of planar polarity
Sox100B	100B2	26,891,079..26,909,129	sequence-specific DNA binding transcription factor activity; DNA bending activity (predicted)	male gonad development
CG11317	100B2	26,916,429..26,923,040	zinc ion binding (predicted)	neurogenesis
CG15553	100B2	26,924,590..26,926,891	transmembrane transporter activity (predicted)	transmembrane transport (predicted)
CG11318	100B2	26,932,088..26,934,910	G-protein coupled receptor activity (predicted)	G-protein coupled receptor protein signaling pathway (predicted)

Prosalpha3T	100B3	26,947,855..26,949,239	endopeptidase activity	ubiquitin-dependent protein catabolic process (predicted)
CG15554	100B3	26,951,873..26,953,058	unknown	unknown
CG15556	100B3	26,953,687..26,956,388	G-protein coupled receptor activity (predicted)	G-protein coupled receptor protein signaling pathway (predicted)
CR43458	100B4	26,958,052..26,958,772	unknown	unknown
Gycbeta100B	100B4	26,960,114..26,997,242	guanylate cyclase activity	cGMP biosynthetic process
stops	100B4-5	26,983,878..26,994,825	unknown	deactivation of rhodopsin mediated signaling
CG31004	100B5	27,004,111..27,021,484	unknown	cell-matrix adhesion (predicted)
bnk	100B5	27,018,946..27,020,517	actin binding (predicted)	actin filament organization; cellular membrane organization; cytokinesis, actomyosin contractile ring assembly; cellularization
CG1544	100B5	27,022,640..27,029,450	oxoglutarate dehydrogenase (succinyl-transferring) activity (predicted)	tricarboxylic acid cycle (predicted)
chp	100B5-6	27,029,904..27,036,452	unknown	rhabdomere development; homophilic cell adhesion
zwilch	100B6	27,036,685..27,039,072	unknown	mitotic cell cycle spindle assembly checkpoint; mitosis
CG1542	100B6-7	27,039,170..27,040,506	unknown	rRNA processing (predicted)
CG15561	100B7	27,040,530..27,041,628	nucleic acid binding (predicted)	unknown
CG1746	100B7	27,041,867..27,045,352	hydrogen-exporting ATPase activity, phosphorylative mechanism (predicted)	proton transport (predicted)
RNaseP:RNA	100B7	27,043,086..27,043,384	ribonuclease P activity	unknown

mir-4949	100B7	27,044,996..27,045,094	unknown	unknown
CG12054	100B8	27,046,426..27,056,645	zinc ion binding; nucleic acid binding (predicted)	unknown
CG15563	100B8	27,057,407..27,058,058	unknown	unknown
CG15564	100B8	27,058,342..27,061,003	unknown	unknown
qless	100B8	27,061,751..27,078,326	trans-hexaprenyltranstransferase activity (predicted)	negative regulation of apoptosis; cellular response to stress; neuroblast development
mRpL32	100B8	27,078,624..27,079,408	structural constituent of ribosome (predicted)	translation
CG1750	100B8	27,079,373..27,080,622	methionyl-tRNA formyltransferase activity (predicted)	biosynthetic process; conversion of methionyl-tRNA to N-formyl-methionyl-tRNA (predicted)
spn-F	100B8-9	27,080,864..27,082,388	minus-end-directed microtubule motor activity	oocyte axis specification; oocyte microtubule cytoskeleton organization; bristle morphogenesis; regulation of cytoskeleton organization
CG3669	100B9	27,083,321..27,085,771	carbonate dehydratase activity (predicted)	one-carbon metabolic process (predicted)
CG18672	100B9	27,085,969..27087389	carbonate dehydratase activity; zinc ion binding (predicted)	one-carbon metabolic process (predicted)
CG18673	100B9	27,088,350..27,089,716	carbonate dehydratase activity; zinc ion binding (predicted)	one-carbon metabolic process (predicted)
CG15555	100B9	27,091,452..27,094,233	sodium channel activity (predicted)	sodium ion transport (predicted)
CG11340	100C1	27,128,211..27,130,962	extracellular-glycine-gated chloride channel activity (predicted)	ion transport (predicted)
CG34347	100C1	27,136,896..27,196,498	actin binding; cytoskeletal protein binding (predicted)	unknown

gskt	100C2	27,198,149..27,199,920	protein serine/threonine kinase activity	male gamete generation; male gonad development
CG1607	100C2	27,207,832..27,217,622	amino acid transmembrane transporter activity (predicted)	amino acid transmembrane transport (predicted)
CG31204	100C3	27,217,616..27,219249	unknown	unknown
CG31002	100C3	27,219,304..27,221,020	glucuronosyltransferase activity (predicted)	metabolic process (predicted)
Gcn2	100C3	27,221,380..27,227,735	elongation factor-2 kinase activity (predicted)	positive regulation of cell size; protein phosphorylation
CG11337	100C3	27,227,282..27,231,174	polyribonucleotide nucleotidyl-transferase activity (predicted)	RNA processing; mRNA catabolic process (predicted)
Gprk2	100C3-4	27,230,971..27,283,499	G-protein coupled receptor kinase activity	vitellogenesis; smoothened signaling pathway; regulation of smoothened signaling pathway; defense response to Gram positive bacterium; imaginal disc derived wing vein specification; embryo development; regulation of Toll signaling pathway; regulation of antimicrobial peptide biosynthetic process
lox	100C4	27,284,102..27,285,591	copper ion binding; protein-lysine 6-oxidase activity	cell adhesion; protein modification process (predicted)
CG11334	100C4	27,285,544..27,288,730	translation regulator activity (predicted)	wing disc development
CG11333	100C4	27,289,067..27,289,846	catalytic activity (predicted)	metabolic process (predicted)
mey	100C4	27,324,961..27,332,716	unknown	regulation of embryonic cell shape
nyo	100C4-6	27,368,754..27,391,887	unknown	regulation of embryonic cell shape

gammaCop	100C6	27,397,865..27,401,429	binding; structural molecule activity (predicted)	open tracheal system development; biological regulation; system development; cellular process; cellular component organization or biogenesis; localization; multicellular organismal development; regulation of tube architecture, open tracheal system; gland morphogenesis; chitin based cuticle development
pygo	100C6	27,401,557..27,406,412	protein binding	biological regulation; cellular process; system development; multicellular organismal development; post embryonic organ morphogenesis; response to stimulus; instar larval or pupal morphogenesis; cellular component organization or biogenesis; embryonic pattern specification; growth; gene expression