Electronic Thesis and Dissertation Repository

8-22-2017 10:00 AM

Investigating the Role of Fruitless in Behavioural Isolation Between Drosophila melanogaster and Drosophila simulans

Jalina Bielaska Da Silva, *The Unviersity of Western Ontario*

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Ecology and Evolutionary Biology Commons

Recommended Citation

Abstract

Behavioural isolation is a prezygotic mechanism that is usually determined by female preference, such as seen with the rejection behaviour exhibited by *Drosophila simulans* females to *D. melanogaster* males. To confirm the role of a previously identified candidate gene *fruitless* (*fru*) in behavioural isolation, I proposed to disrupt *fru* expression in both *D. melanogaster* and *D. simulans* to allow for the generation of interspecies hybrids expressing only a species-specific allele of *fru*. A reciprocal hemizygosity test would then be used to confirm the role of *fru* in behavioural isolation. Disruptions of *fru* in both *D. melanogaster* and *D. simulans* through the CRISPR/Cas9 system were not achieved, however, a mutation was generated in the *fru* common region exon C4 in *D. melanogaster*. This mutation did not have an effect on mating behaviour, suggesting that the C4 exon in *fru* does not seem to have a role in female mate preference or male courtship.

Keywords

Co-Authorship Statement

I performed all the experimental procedures and drafted the written thesis. Melissa Wong contributed to half of the results to one injection experiment: testing survivability with wire-and-tape dechorionation after injections. Dr. Amanda Moehring made intellectual contributions to experimental design and editorial comments for this thesis.

Acknowledgments

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

a.a.	amino acid
Bal	Balancer chromosome
Bon	Bonus protein
bp	base pair
BTB domain	BR-C, ttk, and bab domain
C1-5	common region exons 1-5 of *fru*
Cas	CRISPR-associated
CHC	cuticular hydrocarbon
chst	*chaste*
CNS	central nervous system
CRISPR	clustered regularly interspaced short palindromic repeats
crRNA	crispr RNA
Df	deficiency
DNA	deoxyribonucleic acid
DSB	double strand break
dsf	*dissatisfaction*
dsx	*doublesex*
F_1	offspring from the crossing amongst the P generation
fru	*fruitless*
Fru^M	male-specific Fruitless protein

fst	*frost*
G_0, G_1	generation 0, 1
GFP	green fluorescent protein
Gr32a	*gustatory receptor 32a*
HDAC1	histone deacetylase 1
HP1a	heterochromatin protein 1a
HR	homologous recombination
NHEJ	non-homologous end-joining
nos	*nanos* promoter
P1-4	promoter exons 1-4 of *fru*
PAM	proto-adjacent motif
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
per	*period*
POZ domain	Pox virus and zinc finger domain
RNA	ribonucleic acid
RE	restriction enzyme
Sb	*Stubble*
sgRNA	single guide RNA
SNP	single nucleotide polymorphism

spin	*spinster*
Sxl	*sex-lethal*
TAE	tris-acetate-EDTA buffer
TBE	tris-borate-EDTA buffer
TALEN	transcription activator-like effector nuclease
Tb	*Tubby*
tra	*transformer*
tra2	*transformer-2*
tRNA	transfer RNA
tracrRNA	trans-activating crRNA
w	*white*
y	*yellow*
ZF	zinc-finger
ZFN	zinc-finger nuclease

1   Introduction

1.1   Speciation and behavioural isolation

The evolutionary process whereby new distinct species arise from a common ancestor is known as speciation. There are numerous definitions for how to differentiate species, but most commonly the Biological Species Concept is used, which states that species can be regarded as groups of interbreeding natural populations that are reproductively isolated from other such groups (Mayr, 1942; Coyne and Orr, 2004). The development of reproductive isolation barriers, regarded as biological features of organisms, prevents gene flow between individuals of different species, maintaining species diversity (Dobzhansky, 1937). These reproductive isolation mechanisms can be divided into prezygotic and postzygotic isolating barriers (Dobzhansky, 1937). Postzygotic isolating barriers are those that occur after the formation of the zygote and pertain to hybrid sterility and inviability. Prezygotic isolating barriers act to impede gene flow before sperm transfer and include behavioural, ecological, and mechanical isolation (Dobzhansky, 1937; Coyne and Orr, 2004). While prezygotic barriers can act on their own as a species barrier, behavioural isolation experiences stronger selective pressure in the presence of postzygotic barriers (Coyne and Orr, 1989; Liou and Price, 1994). When hybrid offspring would be unfit, individuals can increase their reproductive fitness by mating with their own species and rejecting those from another species that are not suitable.

For species that are in contact, prezygotic barriers are the strongest barriers as they reduce gene flow proportionally more than postzygotic barriers (Dobzhansky, 1937; Coyne and Orr, 2004). Within prezygotic isolation, behavioural isolation is often regarded as one of the most important impediments to gene flow between species and is possibly one of the first barriers to initiate speciation (Coyne and Orr, 2004). Behavioural isolation consists of behavioural differences in courting and mating signals that prevent different species from mating (Dobzhansky, 1937). The signal from individuals of one sex will elicit a preference in individuals of the opposite sex from the same species, but not from a different species, serving as species-specific cues to avoid heterospecific

mating. This incompatibility of mating signals may often be displayed with divergence in male courtship behaviour (reviewed in: Spieth, 1974; Ritchie *et al*., 1999). If a male's signal is ineffective or repellent to females, she will show a lack of preference for that male and reject his attempt to mate.

Courtship behaviour involves a number of different signals including auditory, visual, tactile, and chemical signals. In Moltoni's warblers, *Sylvia cantillans*, auditory signals play a role in courtship. Different populations of this species distinguish mating songs between other subspecies, and responded most strongly to songs of their own population (Brambilla *et al*., 2008), demonstrating behavioural reproductive isolation between populations within this species. The butterfly species *Pieris occidentalis* and *P. protodice* rely on visual signals when courting in order to discriminate heterospecific mates (Wiernasz and Kingsolver, 1992). Female *P. occidentalis* use the male dorsal forewing melanin pattern to discriminate against *P. protodice* males, insuring that they only mate with conspecific males (Wiernasz and Kingsolver, 1992). In grasshoppers, both auditory and chemical signals are used in courtship; between two closely related species, *Chorthippus biguttulus* and *C. mollis*, males use species-specific calling songs, as well as sex-specific information from female cuticular hydrocarbons (CHCs), for sex recognition and mate attraction (Finck *et al.,* 2016). These signals elicit a courtship song from the males towards conspecific females, but not heterospecific females (Finck *et al.,* 2016). Similarly, many species within the genus *Drosophila* (commonly referred to as the fruit fly) also rely on species-specific courtship songs and CHCs to discriminate against heterospecific mates.

Like all traits, species-specific behaviour leading to female and male mate preference can be linked back to genetic variation. It is known that there is genetic variation to account for preference, as mate recognition systems change quickly, which we see with the high occurrence of prezygotically isolated sibling species that have little morphological divergence (reviewed in: Butlin and Ritchie, 1989). Further, the genetic basis of behavioural isolation can vary depending on the strain being examined, as some populations of species do not manifest the same discrimination genes, making some genes strain-specific as opposed to species-specific (Carracedo *et al*., 2000; Moehring

lab, unpublished data). Genetic variations that have been linked to species-specific behaviour have been associated mostly with male sexual traits, such as those that control male song. In the Hawaiian cricket, the variation of pulse rate in courtship song amongst *Laupala paranigra* and *L. kohalensis* has been attributed to the variable expression of a number of genes that each have a small effect on pulse rate (Shaw, 1996). The courtship song of male hybrids between the grasshoppers *Chorthippus albomarginatus* and *C. oschei* showed higher song variation than the variation that exists between the parents. This change in song can be attributed to an additive-dominance effect of the parental genes in the hybrid, indicating a genetic basis to courtship song (Vedenina *et al*., 2007). One of the first studies to directly link mate preference to genes was done with the ladybird *Adalia bipunctata* where females of different populations showed preference for specific colours and patterns of males (Majerus *et al*., 1982). Females preferred the melanic patterned males, with the melanic form being attributed to a dominant locus (Majerus *et al*., 1982).

The evolution of genes contributing to male and female mating preference behaviour are likely to have arisen in one of two ways: originating as one common gene with pleiotropic effects in males and females, or two different sex-specific genes that have co-evolved together (reviewed in: Butlin and Ritchie, 1989). Yet little is still known about the nature through which sexual selection acts on genes to give rise to male and female preference. In *Drosophila*, genes affecting behavioural isolation between different species have been mapped to chromosomes or small genomic regions (Zouros, 1981; Ting *et al*., 2001; Gleason and Ritchie, 2004; Moehring *et al*., 2006; Laturney and Moehring, 2012a). It is unknown whether within- and between-species mating preference may have the same genetic locus, as proposed by Carracedo *et al* (1989). One study has found chromosomal regions affect species-specific mate preference did not overlap with regions for conspecific preference, showing that there could be different genes contributing to the behaviours of within vs. between species (Gleason and Ritchie, 2004).

1.2 *Drosophila* as a model for behavioural isolation

To study the genetic basis of behaviour, *D. melanogaster* is an ideal model organism. *Drosophila* have a short generation time, are easy and cheap to maintain, and are easy to collect in large quantities, which is important to generate large samples sizes in behaviour studies where there is high variation in a trait. Most importantly, they have a variety of genetic information and tools available, as well as a repertoire of stereotypical behaviours that they exhibit, such as courtship behaviour, making it possible to study how different genes effect behaviour. *D. melanogaster* can even be crossed in the lab with related species for the purpose of mapping species-specific mating behaviours to regions of their genomes (reviewed in: Laturney and Moehring, 2012b).

The single-gene mutant approach in *D. melanogaster* has allowed scientists to mutate one gene at a time to see the underlying effect this mutation has on behaviours (reviewed in: Sokolowski, 2001). For example, the discovery that the *period* (*per*) gene is an underlying genetic component to eclosion rhythm and thus circadian rhythm in flies was one of the first to link a gene to complex behaviours (Konopka and Benzer, 1971; reviewed in: Sokolowski, 2001; reviewed in: Panda *et al*., 2002). Discovering the mechanisms underlying circadian rhythm has made it possible to map circadian rhythm to lateral neurons in brain regions of *D. melanogaster*. This has allowed for the mechanism to be identified in other invertebrate and vertebrates species and for *D. melanogaster* to be a model of circadian rhythm-linked behaviours such as psychiatric diseases in humans (reviewed in: Sokolowski, 2001; reviewed in: Panda *et al*., 2002; reviewed in: Zordon and Sandrelli, 2015). Thus *D. melanogaster* is also a model for neural circuitry, allowing for *D. melanogaster* to be used to study learning, memory, and behaviour. This provides the opportunity to map candidate genes of behavioural isolation to neuronal pathways in *Drosophila*. In summary, the ability to distinguish courtship behaviour in *Drosophila* to identify different behaviours that contribute to mate preference, and finally to use the genetic tools available in *Drosophila*, partnered with behavioural tests, to isolate candidate genes are what makes *Drosophila* a valuable model for behavioural isolation.

1.2.1 *Drosophila* courtship and mating behaviour

To attract a conspecific female, a male of the *D. melanogaster* subgroup performs a stereotypical courtship ritual: (i) he orients himself towards the female; (ii) taps her on the abdomen with both of his forelegs; (iii) follows her and vibrates his wing to produce a species-specific courtship song; (iv) licks the female's extended genitalia by extending his proboscis; and (v) curls his abdomen and attempts to copulate (reviewed in: Hall, 1994; reviewed in: Greenspan and Ferveur, 2000). In response to the male's behaviour, if the female is receptive to his mating signals, she will slow down her locomotive activity and adjust her external genitalia to make them accessible to the male. However, if the female is not receptive to the male, she will perform rejection behaviours by flicking her wings, kicking her legs, or extruding her ovipositor (reviewed in: Hall, 1994).

In *Drosophila*, the courtship behaviour can differ between species and within species subgroups (reviewed in: Spieth, 1974; Cobb *et al.*, 1985). For example, different species will exhibit different patterns during courtship song. *D. virilis* males only extend their wings 10-14° and vibrate in small amplitudes of displacement up and down. Conversely, *D. melanogaster* males extend their wings 90° and vibrate up and down at a larger displacement while *D. planitibia* extend their wings 160° and move their wings backwards and forwards as well as up and down (reviewed in: Spieth, 1974).

Behaviours also vary among closely related species within subgroups. For example, in the *D. melanogaster* subgroup, both *D. melanogaster* and *D. simulans* produce a pulse and sine song in their courtship song, but only *D. melanogaster* males change the duration of each song type as they mature. *D. simulans* males will keep their songs the same no matter the age of the male (Moulin *et al.*, 2001). Additionally, *D. simulans* males show scissoring wing vibration behaviour during their song, which is not done by *D. melanogaster* males (Cobb *et al.*, 1985).

The pulse song is species-specific, allowing for females to recognize conspecific males, while the sine song is important for female sexual stimulation (von Schilcher, 1976). These components of the courtship song are one factor that helps *Drosophila* females to identify a suitable conspecific mate; if the male is heterospecific the female will reject the

male, preventing gene flow between different species. Additionally, the pheromones of *Drosophila* (CHCs) are species-specific, allowing for females to recognize conspecific mates when undergoing courtship. For example, *D. serrate* has been found to express the gene *CG3542*, which is associated with the production of methyl-branched CHCs. This gene is not expressed in the closely related *D. birchii* due to sequence variation in the region, meaning *D. birchii* does not produce this form of CHCs, influencing mate choice between the species (Chung *et al*., 2014).

1.2.2   Behavioural isolation in *Drosophila*

Behavioural isolation amongst *Drosophila* is often what keeps each of the *Drosophila* species distinct, and is usually dependent on female preference (Coyn and Orr, 2004). In the *D. melanogaster* subgroup, some sibling species demonstrate asymmetrical sexual isolation. This is where males from either species will attempt to mate with heterospecific females, but only females of one species are receptive to heterospecific males. The females from the species that chooses to only mate with conspecific males can be seen as "choosy", but potentially only in the context of this species pair.

The sibling species pair *D. melanogaster* and *D. simulans* are sympatric species inhabiting worldwide regions in Africa, Europe, North America, and South America (Sturtevant, 1920; Lachaise *et al*., 1988; Capy *et al*., 1993). However, *D. melanogaster* and *D. simulans* remain behaviourally reproductively isolated from each other because of asymmetrical sexual isolation (Carracedo *et al*., 2000). The behavioural isolation experienced between these two species arises from the rejection behaviour exhibited by *D. simulans* females; *D. melanogaster* females will mate with *D. simulans* males (albeit at reduced frequency), but *D. simulans* females will rarely mate with *D. melanogaster* males (Figure 1; Watanabe and Kawanishi, 1979; Moulin *et al*., 2004; reviewed in: Nanda and Singh, 2012). Sterile *D. melanogaster/D. simulans* hybrid females, generated through crossing *D. melanogaster* females with *D. simulans* males, will mate with *D. melanogaster* males (Figure 1). Thus *D. melanogaster* female receptivity behaviour is dominant or semi-dominant to *D. simulans* female rejection behaviour as the female chooses to mate with a *D. melanogaster* male, behaving like a *D. melanogaster* female

instead of a *D. simulans* female (Carracedo *et al.*, 2000; reviewed in: Laturney and Moehring, 2012a; Laturney and Moehring, 2012b).

As characteristic mating behaviours can be identified between *D. melanogaster* and *D. simulans*, then there is likely a genetic basis for these behaviours. This means genes involved in courtship and mating can be the same genes that cause behavioural isolation and thus speciation. At the moment, some genes have been identified to account for species-specific behaviour, such as *per*, a gene known for its involvement in circadian rhythm (Konopka and Benzer, 1971), and who's genetic function has also been associated with species-specific courtship song in *D. melanogaster* and *D. simulans* (Wheeler *et al.*, 1991). This same gene affects female preference behaviour, whereby females of different species show preferences for different times of mating, as associated with *per* expression (Tauber *et al.*, 2003). However, most genes associated with courtship behaviour have only been associated with female conspecific mating preference, not heterospecific discrimination. For example, *dissatisfaction* (*dsf*) is a gene that acts in a subset of female neurons such that when it is mutated, it causes females to resist conspecific male courtship and take longer to mate (Finley *et al.*, 1997; Finley *et al.*, 1998; O'Kane and Asztalos, 1999). The gene *chaste* (*chst*) also makes females less sexually receptive to conspecific males by causing strong mate refusal in females with a *chst* disruption mutation (Juni and Yamamoto, 2009). Females with a mutation in *spinster* (*spin*) also show subnormal receptivity to copulation by exhibiting rejection behaviours such as decamping, kicking, and fending when they were courted by conspecific males (Suzuki, *et al.*, 1997). There is still no clear understanding of how many genes can affect female behaviour, especially female species-specific preference, nor how their genetic expression can cause different female behaviours in different species. Female preference behaviour is an important contributor to maintaining reproductive isolation between species as it reinforces the divergence of two separate species by preventing gene flow (Coyne and Orr, 2004). By understanding what genes underlie female behaviour, we can uncover genetic variants, such as mutations or alterations in genes, which cause heterospecific differences in mating behaviour. Candidate genes for female preference can then be confirmed through gene disruptions and behavioural tests (see below), followed by mapping of the neural pathway upon which genes linked to mating

behaviour act in order to find where along the path signals differ to give rise to different mating behaviours.
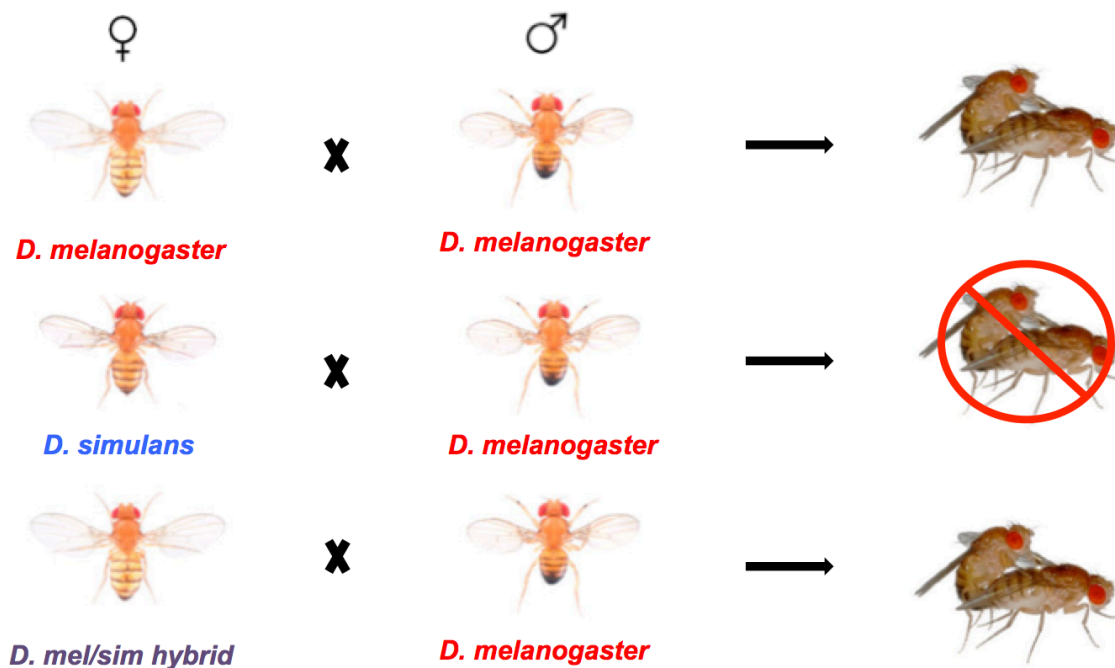
Figure 1: Schematic of the behavioural isolation that occurs between *D. melanogaster* and *D. simulans.*

Female and males from the same species (i.e. female *D. melanogaster* and male *D. melanogaster*) will mate with each other. However, a *D. simulans* female does not mate with a *D. melanogaster* male. *D. melanogaster/D. simulans* hybrid females mate with *D. melanogaster* males, albeit at reduced frequencies, exhibiting behaviour more similar to female *D. melanogaster* than *D. simulans* rejection. Fly images were created from photographs taken from FlyBase (2016).

1.2.3   Commonly-used genetic tools in *Drosophila*

Drosophila became a prominent model to study genetics back in 1910 when Thomas Hunt Morgan used *D. melanogaster* to define genes and show that genes are found on chromosomes (Kohler, 1994; reviewed in: Jennings, 2011). *D. melanogaster* was one of the first model organisms to have genetic tools to study gene expression (reviewed in: Arias, 2008), and continues to be on the leading edge of technological advances, making it a popular model organism for genetics. It was one of the first organisms to have its genome sequenced: Craig Venter's team released the first sequence of the *D. melanogaster* genome in March 2000 (Adams *et al*., 2000). The sequence and annotation of the genome became publicly available on FlyBase (Tweedie *et al.,* 2008), where the genomic sequences of other *Drosophila* species are also available, notably the full *D. simulans* genomic sequence (Hu *et al*., 2013).

The wide diversity of techniques available in *D. melanogaster* allow for the manipulation of genes to study their phenotypic effects. Two types of tools exclusive to *Drosophila* are outlined in the sections below. Mutations in nearly every gene have been generated and lines are readily available for order through stock centres such as the Bloomington *Drosophila* Stock Center, along with targeted expression lines for the use of specific tools such as GAL4/UAS, RNAi, FLP/FRT, and CRISPR/Cas9. However, the genetic tools available for use in *Drosophila* are mostly restricted to *D. melanogaster*. With the advancement of such tools like CRISPR, this limitation has changed, giving us a better ability to study what genes are affecting behaviour in *D. simulans*.

1.2.1.1   *P*-elements

P-elements are a type of mobile DNA element that was introduced into the *D. melanogaster* population around 1950 through horizontal gene transfer with another *Drosophila* species (reviewed in: Ryder and Russell, 2003). *P*-elements have also recently been discovered in *D. simulans* populations (Kofler *et al*., 2015).

The first use of *P*-elements as a transgenic tool was done by Spradling and Rubin (1982), who injected *P*-elements into *Drosophila* embryos and were able to recover flies that had intact copies of the *P*-elements inserted into genes. *P*-elements are a popular genetic tool

and have been used to disrupt gene expression for approximately 65% of all *D. melanogaster* genes. Researchers can readily obtain 95% of these transgenic *P*-element insertion fly lines from the Bloomington *Drosophila* Stock Center (reviewed in: Venken and Bellen, 2005).

However, as Spradling and Rubin (1982) discovered, the *P*-element insertions are not randomly inserted into the genome, appearing at a wide variety of chromosomal sites, but *P*-elements are unable to target specific genes. This is one of the drawbacks for using *P*-elements for gene tagging and gene disruptions of specific targets. Additionally, *P*-elements also exhibit insertion bias, which can be another problem if specific targets are needed. This is due to the occurrence of hot spots and cold spots for *P*-element insertions in the genome, preventing certain genes from ever being targeted (reviewed in: Venken and Bellen, 2005).

1.2.1.2   Balancer chromosomes

The use of balanced lethals, known as balancer chromosomes, in *Drosophila*, has provided researchers with an easy way to maintain mutations in stocks (e.g. gene disruptions) (Arias, 2008). Balancer chromosomes help to maintain homozygous lethal mutations in populations such that linkage between alleles can be maintained and prevent recombination amongst multiple alleles on the same chromosome (Bloomington *Drosophila* Stock Center). To achieve this, all balancer chromosomes have two main features. First, they must contain recessive deleterious mutations that cause lethality or sterility when homozygous, ensuring that the non-balancer chromosome must always be present (Bloomington *Drosophila* Stock Center). The second feature, inversion breakpoints, prevents the recovery of recombinant chromosomes by inhibiting synapsis and produces lethal aneuploid gametes when single crossovers within inversions do occur (Bloomington *Drosophila* Stock Center). Additionally, many balancers have genetic elements associated with them, making it easier to score phenotypically the inheritance of a balancer in a population. One common feature is a dominant visible marker, which provides a clear means of determining if individuals inherited the balancer chromosome or non-balancer chromosome. An example of one of these markers is the stubble hair phenotype associated with the *TM3* balancer, which contains a mutation in the *Stubble*

(*Sb*) gene. Though balancers are mostly restricted to *D. melanogaster*, some strains are available for other species, which contain an inversion that can be used in a similar manner to the balancer chromosome.

1.2.4   Deficiency mapping

There is a limit to using *Drosophila* to identify candidate genes for behavioural isolation: crossing males and females of different species is usually difficult because the crosses do not produce a lot of offspring, the offspring are unhealthy, or only sterile hybrid offspring are produced. *D. melanogaster*, the species with the most genetic tools, does not produce fertile F1 offspring with any of its sibling species, making it impossible to employ the traditional recombinant mapping approach to narrow down candidate genes for behavioural isolation.

Deficiency mapping can be employed to overcome the limitations of not being able to use recombinant mapping to identify candidate regions of behavioural isolation in *D. melanogaster* sibling species. Deficiency mapping relies on using the genetic tools of transposable elements (commonly *P*-elements) or FLP/FRT (Parks *et al*., 2004), and balancer chromosomes to identify genes that contribute to quantitative traits such as mating behaviour (Moehring and Mackay, 2004) and longevity (Pasyukova *et al*., 2000). This technique is even effective in mapping a trait down to a single gene (Pasyukova *et al*., 2000; Moehring and Mackay, 2004). Deficiency mapping requires the use of deficiency fly stocks, whereby the lines are entirely diploid except for a single region that is hemizygous due to a deletion on one of the homologous chromosomes. The deficiencies are maintained over a balancer chromosome, allowing for a visible marker to score for individuals with the balancer and maintain the deficiency in the lines, as deficiencies are usually homozygous inviable. To use deficiency mapping to identify candidate regions for behavioural isolation, a recessive gene that controls a trait of interest can be crossed to the deficiency line. For example, when looking at female species-specific mate preference between *D. melanogaster* and *D. simulans*, female *D. simulans* rejection behaviour of *D. melanogaster* males is recessive to *D. melanogaster* female receptive behaviour. Therefore, by crossing wild-type *D. simulans* to *D. melanogaster* bearing a deficiency, the female hybrids can have one of two genotypes

(Figure 2A). If females inherit the balancer chromosome from her deficient father, all of her recessive *D. simulans* genes will be masked by *D. melanogaster* genes in the chromosomal region covered by the balancer. Alternatively, if the female inherits the deficiency, any recessive *D. simulans* genes in that region will be unmasked and thus expressed, as there is no *D. melanogaster* allele present. If the females that have the deficiency now act like *D. simulans* females and reject *D. melanogaster* males, then the *D. simulans* alleles within that region are candidate genes for female preference behaviour. To control for genetic background and hemizygosity effects, *D. melanogaster* with the deficiency are crossed wtih wild-type *D. melanogaster* to make two genotypes again: *D. melanogaster* with the balancer or *D. melanogaster* with the deficiency (Figure 2B). This is done to ensure that the effect of *D. simulans* behaviour in hybrids is not due to hemizygosity or an extraneous factor elsewhere in the genome, but the behaviour is really due to unmasking the *D. simulans* alleles (Figure 2).

Previous work used deficiency mapping to identify the genetic basis of *D. simulans* rejection behaviour towards *D. melanogaster* (Laturney and Moehring, 2012b). Five regions on the right arm of the third chromosome were identified that contain genes potentially linked to female rejection behaviour underlying behavioural isolation in this species pair (Laturney and Moehring, 2012b). Additional fine-mapping and tests of individual genes identified *fruitless* (*fru*), a gene that is involved in the sex-determination pathway in *Drosophila*, as a candidate gene that contributes to female rejection behaviour (Moehring lab, unpublished data). In hybrid females where the *D. melanogaster* allele of *fru* is knocked out and only the *D. simulans* allele is being expressed, females have shown decreased mating with *D. melanogaster* males, thus acting like *D. simulans* females, when compared to hybrids that are also expressing *D. melanogaster fru* (Moehring lab, unpublished data).
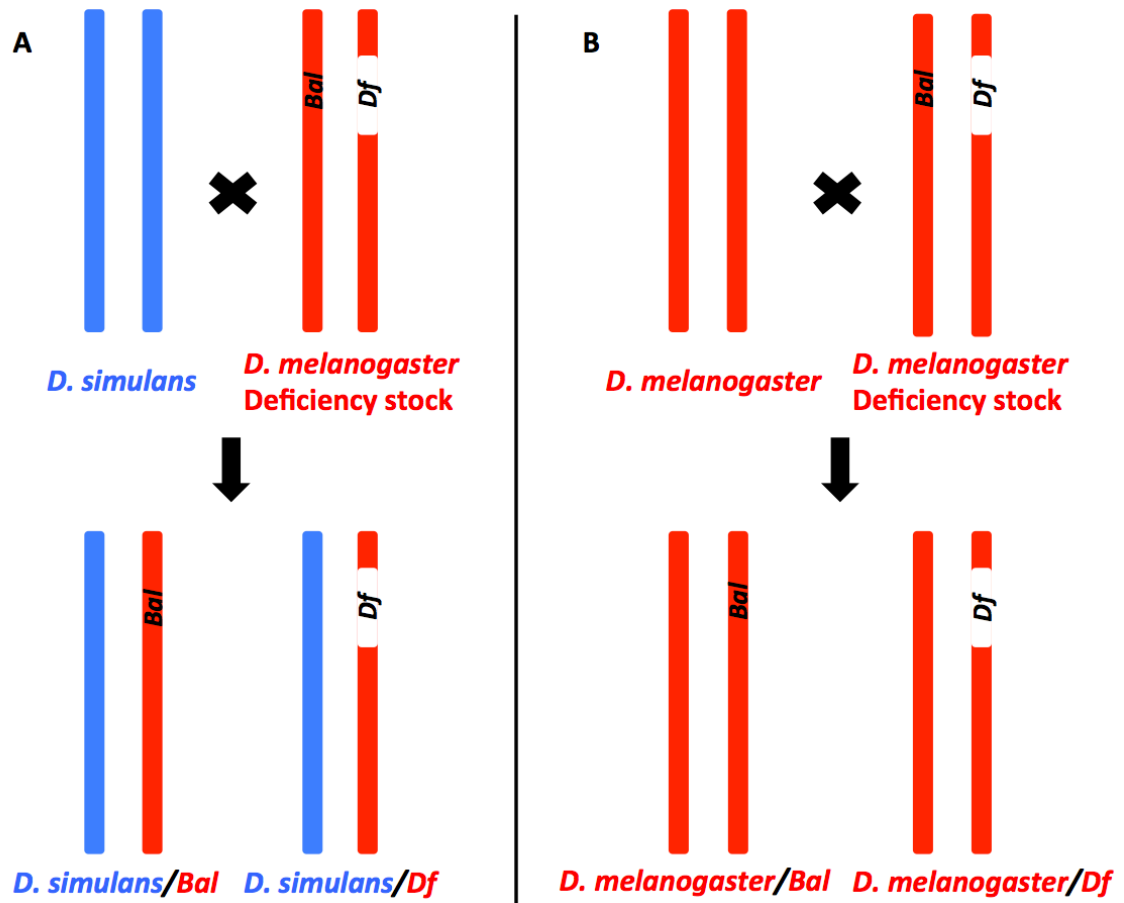
Figure 2: Crosses used to test the four genotypes used in deficiency mapping.

A. A heterospecific cross between *D. simulans* (blue) and *D. melanogaster* (red) deficiency stock to create two hybrid females, one with the balancer (*Bal*) and one with the deficiency (*Df*). The deficiency hybrid will allow for the unmasking of *D. simulans* alleles in that region of the chromosome. **B.** A conspecific cross between *D. melanogaster* to generate two control females. Controls are used to ensure the *D. simulans* recessive behaviour being unmasked is not due to the effect of having a single allele expressed. Bars represent homologous chromosomes with the broken chromosome representing a deficiency.

1.3. The *fruitless* gene in *D. melanogaster*

1.3.1   *fruitless* in the sex determination pathway

The *fru* gene is involved in the sex determination pathway in many insect species, from the mosquito *Anopheles* to *Drosophila* (Gailey *et al.,* 2006). The sex determination cascade that results in sexual differentiation in *Drosophila* takes on the key strategies observed in other sex determination pathways in insects, namely a primary genetic signal that is different in males and females due to the use of a switch gene towards the end of the pathway (Schutt and Nothiger, 2000). Male or female expression is first induced by the ratio of X chromosomes to autosomes, leading to sex-specific expression in *Drosophila*. The sex determination cascade (Figure 3) begins with *Sex lethal* (*Sxl*), where sex-specific expression achieved through alternative splicing sees the production of the Sxl protein in females and a lack of protein expression in males due to premature termination of translation (Bell *et al*., 1991). Sxl controls the expression of the downstream gene *transformer* (*tra*); as Sxl is only expressed in females, it splices at the 3' splice site of *tra* allowing for the production of mRNA with an open reading frame (Belote *et al*., 1989). At the same level of that cascade is *transformer2* (*tra2*) where its function is necessary both to prevent male sexual differentiation and to allow for female differentiation (Belote and Baker, 1982; Nagoshi, *et al*., 1988). Together, the expression of these genes regulates the expression of *fru* and the master-switch gene *doublesex* (*dsx*) through alternative splicing, to produce female-specific transcripts (Nagoshi, *et al*., 1988). As far as we know, female-specific *fru* mRNAs are not translated into functional proteins, due to an early stop codon in the sequence, but the female-specific Dsx protein is responsible for inducing female somatic structures and external morphology (reviewed in: Billeter *et al*., 2006a).

Male-specific transcripts of *dsx* and *fru* are both alternatively spliced to produce male-specific proteins, but without the assistance of *tra* or *tra2*; Tra is not expressed in males, allowing for the male-specific Fru protein to be produced by a default splice. In males, the products of *dsx* and *fru* (zinc finger transcription factors) specify different aspects of male differentiation (somatic structures and external morphology) and male sexual

behaviour (Burtis and Baker, 1989; Anand *et al*., 2001; Demir and Dickson, 2005; reviewed in: Billeter *et al*., 2006a; Billeter *et al*., 2006b). The *fru* portion of the sex determination cascade is responsible for nearly all steps of male courtship behaviour with *dsx* likely controlling male song (Ryner *et al*., 1996; Villella *et al*., 1997; Goodwin *et al*., 2000; Baker *et al*., 2001; Kimura *et al*., 2008). Besides fine-mapping and candidate gene behavioural mating tests on *fru* (Moehring lab, unpublished data), there is currently no evidence directly linking *fru* to female species-specific mate preference.
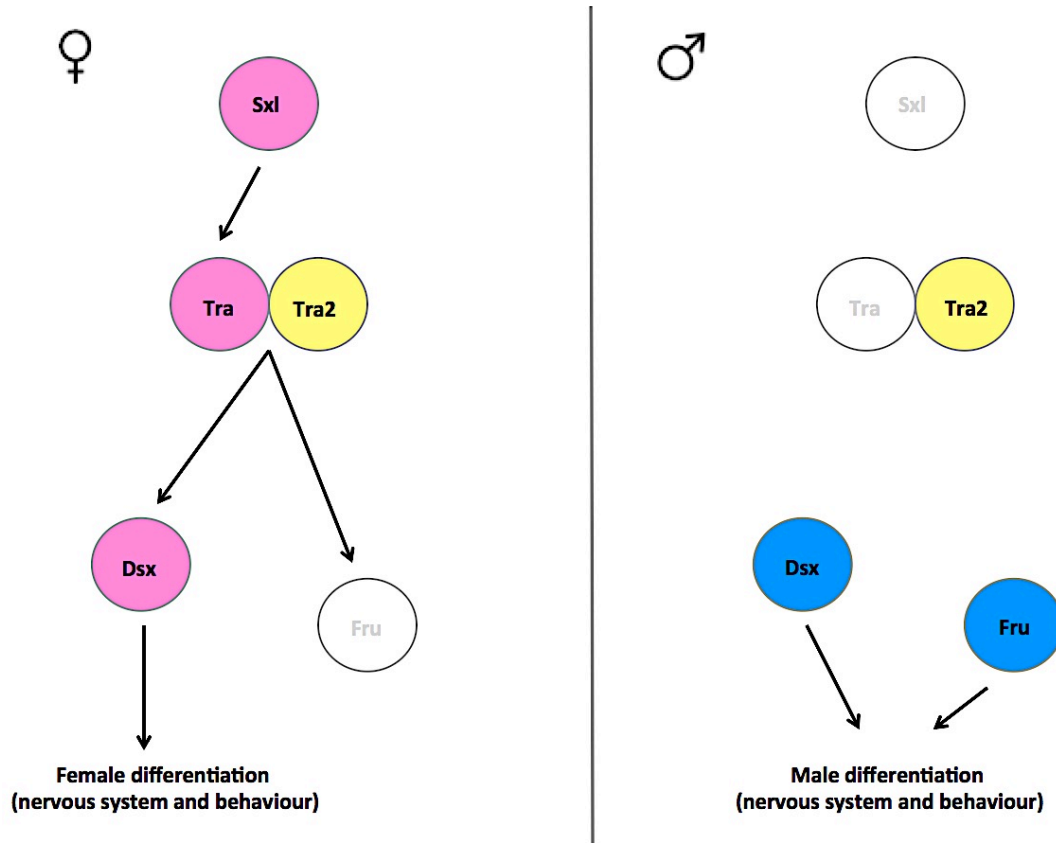
Figure 3: The sex determination pathway in female and male *Drosophila.*

In females (left) Sxl alternatively splices the 3' end of *tra*, which works with Tra2 to alternatively splice *dsx* and *fru* to give female specific products. Only female-specific Dsx, not female-specific Fru, is functional in females, which goes on to specify female differentiation. In males (right), both male-specific copies of Dsx and Fru are expressed, contributing to male differentiation. Fru is responsible for male courtship behaviour. White circles indicate sex-specific proteins that are not expressed in the cascade, with pink indicating female-specific proteins, blue for male-specific proteins, and yellow as non sex-specific. This figure is adapted from Billeter *et al*. (2006a).

1.3.2   *fruitless*: a complex gene

The *fru* gene is an approximately 130 kbp long gene on the third chromosome composed of a number of different alternatively-spliced exons (Figure 4; Ryner *et al*., 1996; Heinrichs *et al.,* 1998; Goodwin *et al.,* 2000; FlyBase: Gramates *et al*., 2017); the P exons (P1-4) at the 5' end, where P1 is sex-specifically spliced into male and female isoforms; common region exons (C1-5), where the common region is found in all *fru* transcripts, are highly conserved between insect species (Clynen *et al.,* 2011), and exons C1 and C2 code for the BTB/POZ domain; and exons A-D at the 3' end that code for alternative zinc-finger (ZF) domains. The 5' P and 3' A-D exons are alternatively spliced, producing transcripts that have a single P exon in combination with one of the A-D exons, with almost all transcripts containing the entire common region (if exon D is in a transcript, C5 is not present) (Figure 4; Ryner *et al*., 1996; Goodwin *et al.,* 2000). Fru is a sequence-specific DNA binding transcription factor, where the BTB domain acts as a protein-protein interaction module and ZF domains act as DNA-binding domains, with a linker region between these two domains (Zollman *et al*., 1994; Ito *et al*., 1996; Ryner *et al*., 1996; Stogios *et al*., 2005).

Transcripts of *fru* associated with P2 to P4 are expressed in both males and females and are essential for development in both sexes, controlling the development of imaginal discs and motoneuronal synapses (Ryner *et al*., 1996; Anand *et al*., 2001). P2 transcripts are expressed in pupae but have the highest expression in the adult central nervous system (CNS) compared to all other transcripts; P3 and P4 transcripts are expressed in early stages of development, as they are integral for the formation of the CNS during embryogenesis (Song *et al*., 2002; Dornan *et al*., 2005; Neville *et al*., 2014).

While P2 to P4 transcripts have common expression in males and females, being expressed in the CNS of both sexes, P1 transcripts have a sex-specific function. In males, P1 transcripts produce male-specific functional proteins (Fru^{M}) while female P1 transcripts are restricted to becoming non-functional truncated proteins due to an early stop codon (Lee *et al*., 2000). If there was no early stop codon in female transcripts, there would be the potential to encode proteins with BTB domains near their terminal end

(encoded by exons C1 and C2) and one of four ZF pairs at their carboxy ends (encoded by exons A-D) (Ryner *et al*., 1996; Goodwin *et al.,* 2000; Anand *et al*., 2001).

It is male-specific expression of P1 transcripts in neurons of the CNS that are responsible for the development of the male-specific abdominal muscle known as the Muscle of Lawrence (MOL) (Gailey *et al*., 1991; Usui-Aoki *et al*., 2000) and male sexual behaviour (Ryner *et al*., 1996; Goodwin *et al*., 2000; Demir and Dickson, 2005). Neurons expressing male-specific Fru protein (Fru^{M}) have been identified in 2% of the male CNS (Usui-Aoki *et al*., 2000) and are believed to be connected together in a circuit that also intersect with olfactory or gustatory neurons, a required pathway in order to exhibit a behavioural response to female sex pheromones (Stockinger *et al*., 2005). However, exactly how Fru^{M} acts at the neuronal level to control male behaviour is still not fully understood.
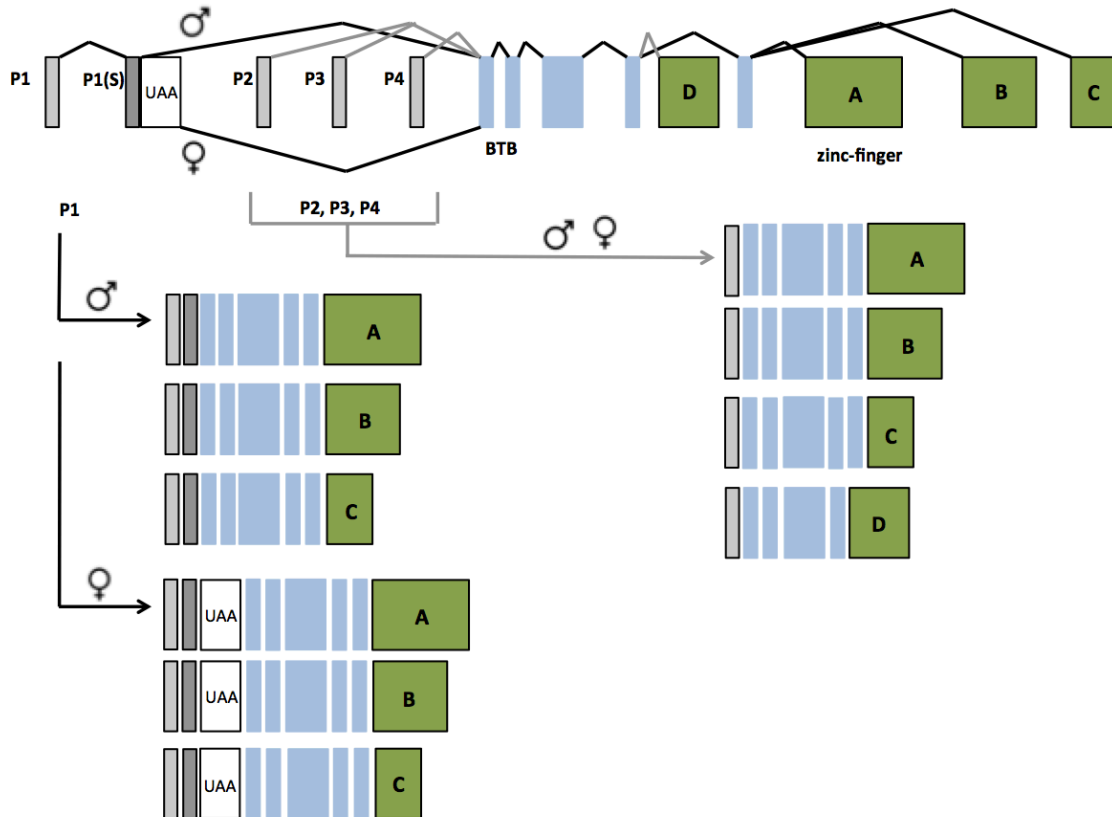
Figure 4: Organization of the *fru* gene and alternative splicing that gives rise to *fru* products in males and females.

The gene *fru* contains four independent first exons (P1-P4) (grey boxes) where the P1 promoter undergoes alternative splicing in males and females under the control of *tra* and *tra2* in the sex determination pathway. It produces three male-specific transcripts that contain the common regions (blue boxes), which makes the BTB domain in Fru protein, and either A, B, or C (green boxes) that compose the zinc-finger domain in the protein. There is no functional protein produced from P1 transcripts in the female. Transcripts from promoters P2-P4 are non-sex-specific (produced in both males and females) and have four alternative products that have the common regions and either A, B, C, or D, composing the final Fru proteins. This figure is adapted from Neville *et al*. (2014).

1.3.3   *fruitless* and behaviour

Fru[M] produces three different proteins, where each has one of three alternative C-terminal zinc-finger DNA binding domains characterized in *fru* as exons A-C (Figure 4). Fru[M] isoforms with A are restricted to a subset of neurons in the male CNS with isoforms B and C more broadly expressed in the male CNS, appearing in most of the cells where Fru[M] expression has been seen (Billeter *et al*., 2006b; Neville *et al*., 2014). Cells forming Fru-specific neurons either express one, two, or all three isoforms (von Philipsborn *et al*., 2014). Isoform A does not seem to be highly involved with male behaviour on its own, but works collectively with the other isoforms of Fru[M] to stimulate male-specific behaviour (Neville *et al*., 2014). Isoforms B and C can work collectively like isoform A, but appear to also have individual roles in male courtship. Both are required for males to exhibit wild-type levels of courtship, whereas isoform B appears to be heavily associated with a male's initiation of courtship towards females, while isoform C has been associated specifically with the production of courtship song and with the formation of the MOL (Billete *et al*., 2006; Neville *et al*., 2014; von Philipsborn *et al*., 2014). Additionally, isoform C seems to be heavily involved in male-specific neuron structure patterns as knocking down expression of this isoform causes female neural patterning (von Philipsborn *et al*., 2014). Eliminating expression of Fru[M] isoform C in a subset of Fru neurons, aSP4 and vAB3, feminizes the neurons morphology (von Philipsborn *et al*., 2014).

Recently, it was found that the Fru[M] protein forms a complex with Bonus (Bon), a transcription cofactor that is expressed in all CNS neurons, including *fru* neurons (Ito *et al*., 2012). The Fru-Bon complex alters chromatin modification to induce gene silencing depending on whether it interacts with HDAC1 or HP1a in order to regulate the level of neural masculinization. When the complex recruits HDAC1, it acts as a positive regulator for masculinization; when it recruits HP1a, it acts as a negative regulator for masculinization. Together, they counteract each other for mediating Fru-dependent sexually dimorphic neuronal development to determine the sexual fate of neurons.

Fru[M] has been found to be part of the same neuronal pathway as Gr32a gustatory neurons, which act in the male foreleg to detect CHCs secreted by other species in order

to recognize when a female is heterospecific and help inhibit heterospecific courtship (Fan *et al.*, 2013). Specifically, Fru^M also functions in subesophageal ganglion aDT6 gustatory neurons, which act to inhibit courtship of females from another species, but these neurons do not directly interact with Gr32a neurons (Fan *et al.*, 2013). This indicates that multiple, distinct mechanisms act in this pathway to control interspecies courtship in males. For females, the Gr32a neurons do not have a role in interspecies mating and so the neuronal basis for female rejection is still unknown (Fan *et al.*, 2013).

The heavy association of three Fru^M isoforms with male courtship through the function of the ZF domains implies that the ZF domains in other transcripts, though not sex-specific, could have a sex-specific effect on the neuronal circuitry. This could imply an association with female behaviour. If lack of Fru^M-specific expression can feminize neurons, then perhaps this, partnered with dimorphic expression of the ZF isoforms of the common Fru proteins, leads to female behaviour. The expression of multiple common isoforms in different subsets of neurons and their interaction with other genes could regulate the female behavioural pathway, just as isoforms of Fru^M assist in the development of neurons, with the assistance of *dsx*, to give rise to male copulatory behaviour (Billeter, *et al.*, 2006b).

1.4   Goal: Confirming the involvement of *fruitless* in female rejection behaviour

As deficiency mapping and candidate gene complementation tests performed in my lab have solidified *fru* as a candidate gene for behavioural isolation (Moehring lab, unpublished data), then *fru* must play a role in female species-specific mate preference behaviour. As the *fru* disruption in *D. simulans* has never been tested before, this study will look to confirm that *fru* does contribute to behavioural isolation between *D. melanogaster* and *D. simulans* by disrupting expression of *fru* in both species. If *fru* does have a role in female species-specific mate preference behaviour, then it is expected that hybrids expressing only the *D. simulans* allele of *fru* will show significantly reduced mating compared to the female hybrids expressing *D. melanogaster fru* allele when these females are paired with *D. melanogaster* males. One way to accomplish this would be to

replace the entire *D. melanogaster fru* gene and its regulatory sequences with the corresponding sequences from *D. simulans* and then score female behaviour. Since the gene is a cumbersome 130 kbp long, and most of the regulatory sequences have not been identified, this approach is currently intractable. Additionally, incorporating the entire DNA region of a large gene like *fru* is very difficult (Cande *et al*., 2014). An alternative approach is to perform a gene disruption test, where the *D. simulans* allele is disrupted, and observe the effect on behaviour. The appropriate behavioural test to use for this aim is the reciprocal hemizygosity test.

1.5 Experimental approach

1.5.1   Reciprocal hemizygosity test with *fruitless* disruption to confirm its role

A reciprocal hemizygosity test uses hybrids that have the same genomes, except at one locus, to test if the gene at that locus has diverged between the two species. This requires crossing two strains in which a gene disruption has been generated (Stern, 2014). In this case, it would be the crossing of *D. simulans* with *fru* disruption to wild-type *D. melanogaster*, and *D. melanogaster* with *fru* disruption to wild-type *D. simulans*. This will allow for the generation of F1 offspring that are genetically identical at all loci except at the location of the mutation, where they only express a species-specific allele (Stern, 2014). In the case of *D. melanogaster* and *D. simulans*, hybrids generated would be a *D. melanogaster/D. simulans* hybrid with the *D. simulans fru* disruption, (only *D. melanogaster fru* expressed), or a *D. melanogaster/D. simulans* hybrid with the *D. melanogaster fru* disruption (only *D. simulans fru* expressed) (Figure 5). The approach chosen to disrupt *fru* in *D. melanogaster* and *D. simulans* in order to carry out a reciprocal hemizygosity test is CRISPR/Cas9 – see below.
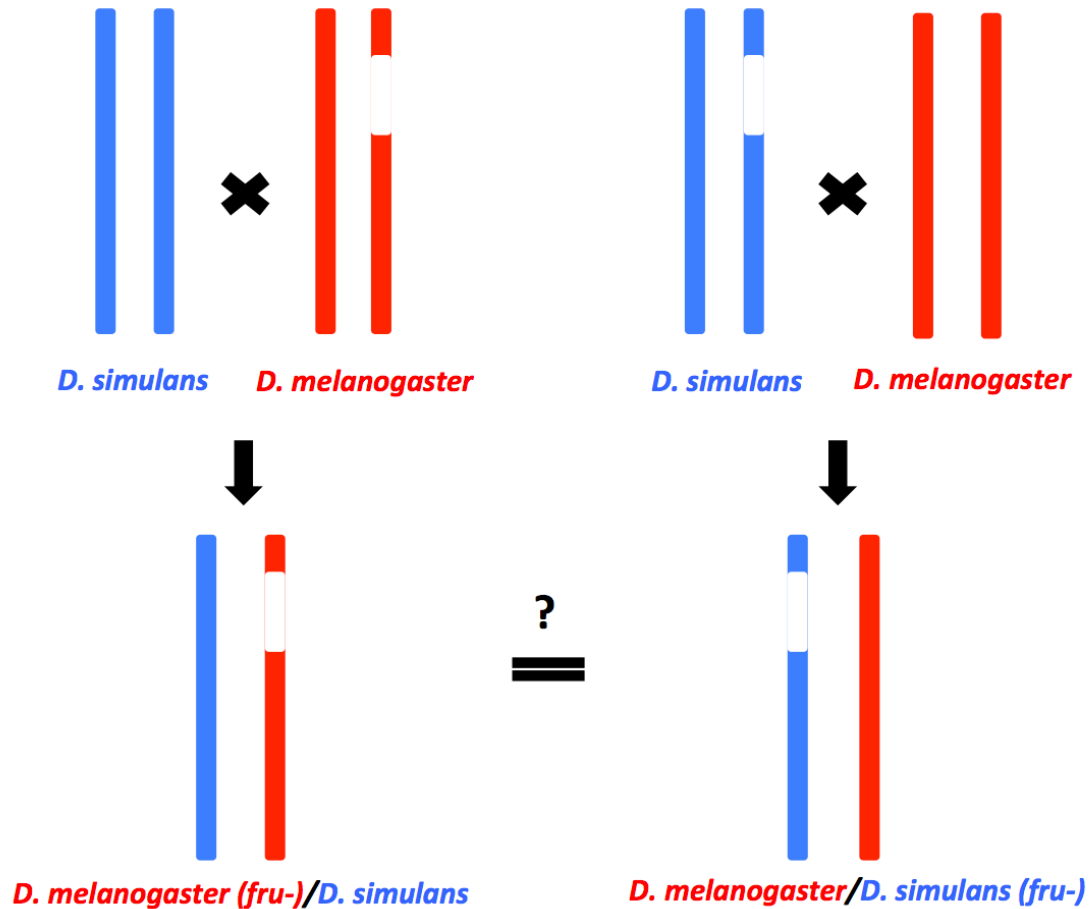
Figure 5: Generation of hybrids with a species-specific allele disruption of *fru* to test behaviour using a reciprocal hemizygosity test.

D. melanogaster (red) and *D. simulans* (blue) strains, two of which were made to have a *fru* knockout (broken lines). When crossed with the opposite species (wild-type), hybrids are generated, both of which have identical genomes except for the one-species *fru* allele that has been knocked out. This test can be used to see if expression of species-specific *fru* alleles in an identical genetic background gives rise to species-specific behaviour. If it does, then the behaviour of the two hybrids would not be equivalent to each other.

1.5.2   CRISPR/Cas9 system

Currently, CRISPR (clustered regularly interspaced short palindromic repeats) is the most efficient genome editing technique for generating double strand breaks (DSB; reviewed in: Beumer and Carroll, 2014). CRISPR has several advantages compared to previously used genome editing techniques (such as ZFNs and TALENs; reviewed in: Beumer and Carroll, 2014). To edit the genome using CRISPR, only a single splicing protein (Cas9) is used, meaning that no protein engineering is required, unlike when using TALENS. CRISPR targeting only depends on base pairing, requiring the production of a simple single guide RNA (sgRNA) to target the sequence of interest. CRISPR also has the ability to produce DSBs at multiple targets (reviewed in: Beumer and Carroll, 2014). CRISPR/Cas9 system is a widely used system for genome engineering in such organisms as yeast (*Saccharomyces cerevisiae*) (DiCarlo *et al.,* 2013; Jakociunas *et al*., 2015), silkworms (*Bombyx mori*) (Wang *et al.,* 2013), mosquitos (*Aedes aegypti*) (Dong *et al*., 2015), beetles (*Tribolium castaneum*) (Gilles, *et al*., 2015) zebrafish (Hwang *et al.,* 2013; Jao *et al.,* 2013), plants and crops (Shan *et al.,* 2013; reviewed in: Belhaj *et al.,* 2013), mice and rats (Li *et al.,* 2013; Wu *et al.,* 2013; Wang *et al.,* 2013), human cells (Ding *et al.,* 2013; Mali *et al*., 2013), and in different *Drosophila* species (personal communication with Best Gene where they expressed they could inject in any species provided to them, but could not guarantee efficiency). Thus, CRISPR/Cas9 can be used in *D. melanogaster* and *D. simulans*, giving it the advantage over tools that are only available in *D. melanogaster* (such as Gal4/UAS) to study species-specific mate preference.

1.5.2.1   History and classification

Originally, the CRISPR system was identified as an adaptive immunodefence system used by bacteria and archaea to fight off invading viruses. The components of the adaptive immune system are broken down into three main stages: acquisition, expression, and interference (reviewed in: Bhaya *et al*., 2011). The first stage begins once a virus attacks, whereby small 32-nucleotide fragments of the viral DNA gets incorporated into the bacterial or archaeal chromosome at the end of a repetitive element, forming the CRISPR loci (Deveau *et al*., 2008; Garneau *et al*, 2010). These CRISPR loci are often

adjacent to *cas* (CRISPR-associated system) genes and have numerous 29-nucleotide palindromic repeats separated by the 32-nucleotide fragments of viral DNA known as spacers (Deveau *et al*., 2008). The incorporation of spacers allows for an adaptive immune system that can recognize virus and cleave the double-stranded DNA that compliments the spacer sequence (Garneau *et al*, 2010). At the second stage, transcription of the repeat and spacer forms CRISPR RNA (crRNA) that base-pairs to trans-activating crRNA (tracrRNA), forming a two-RNA structure that directs Cas9 endonuclease to the invading virus. This begins the third stage where the crRNA and tracrRNA with Cas9 initiate DSBs in the target DNA sequence, interfering with virus replication and imparting immunity to the host (Jinek, *et al.,* 2012, reviewed in: Bhaya *et al*., 2011).

The CRISPR/Cas system has been classified into three main groups known as type I, type II, and type III (Jansen *et al*., 2002; Makarova *et al*., 2011; Sinkunas *et al*., 2011; reviewed in: Bhaya *et al*., 2011; Jinek *et al.,* 2012) that all have a large number of diverse *cas* genes, but with *cas1* and *cas2* being common to all systems (Haft *et al.,* 2005, reviewed in: Bhaya *et al*., 2011). The specific nuclease genes can be classified as universal (*cas1* and *cas2*), type-dependent (*cas4, cas5, cas6, cas7,* and *cas8*), and signature (*cas3, cas9,* and *cas10*) (Haft *et al.,* 2005; reviewed in: Bhaya *et al*., 2011). Type-dependent genes are often associated with expression and/or interference while signature genes are usually only associated with interference (reviewed in: Bhaya *et al*., 2011). The expression of these genes is what helps to classify different CRISPR systems into the three types.

Type I is the most diverse of the three CRISPR types, containing six subtypes (type I-A through to type I-F) that all cleave DNA with assistance from the exclusive Cas3 endonuclease (Makarova *et al*., 2011; Sinkunas *et al*., 2011). Type III has two variations, type III-A and III-B, based on the functional differences between the two: type III-A targets plasmid DNA *in vivo,* as seen in the immune system of *Staphylococcus epidermidis* while type III-B, found in *Pyrococcus furiosus,* only cleaves single-stranded RNA substrates *in vitro* (Marraffini and Sontheimer, 2008; Hale *et al.,* 2009; reviewed in: Wiedenheft *et al*., 2012). The finding of type III-B CRISPR shows the complexity and

mechanistic differences that exists within the CRISPR systems, as it is one of the first subtypes to cleave RNA instead of DNA (Hale *et al.,* 2009).

The type II CRISPR/Cas9 system is composed of a guide RNA (gRNA) with two parts (crisprRNA and tracrRNA) that bind the DNA target sequence and Cas9 endonuclease. Cas9 then creates a DSB adjacent to the target sequence (Figure 6A; Jansen *et al*., 2002; Jinek *et al.,* 2012). The Cas9 protein requires a 20-nucleotide guide sequence complementary to the target sequence (gRNA; specifically the crRNA component) and a conserved protospacer adjacent motif (PAM) sequence, composed of nucleotides NGG, downstream of the target sequence (Jinek *et al.,* 2012). NGG occurs approximately every 8 bases in the DNA, making it easy to acquire targets. Additionally, the requirement of having NGG at the 3' end of the target sequence can be relaxed to also include NAG, allowing for more potential targets (Mali *et al.,* 2013).

1.5.2.2   CRISPR/Cas9 mechanism: generating targeted DSBs

The type II system has been simplified to a two-component system for use *in vivo*; it is composed of Cas9 and a single synthetic guide RNA (sgRNA) that contains the minimal function sequence from the gRNA (Figure 6B; Mali *et al.,* 2013; reviewed in: Bassett and Liu, 2013; Bassett, *et al*., 2013; Gratz, *et al*., 2013; Ren, *et* al., 2013; Bassett and Liu, 2014; Gratz, *et al*., 2014; Ren *et al.,* 2014). One sgRNA can be used to create a DSB, paired sgRNAs can be used to target complimentary strands for higher specificity, or two sgRNAs aimed at two different sequence targets in a gene can be used to delete the intervening sequence. This latter scenario can also be paired with a single stranded oligonucleotide donor sequence to integrate short sequences at the cleavage site (Gratz *et al*., 2013; reviewed in: Bassett and Liu, 2013; Ren *et al*., 2013). In addition, transgenic flies expressing the Cas9 protein have been produced to optimize the system, allowing for only sgRNA to need to be injected into the embryo. Cas9 expression in *Drosophila* can be driven by either the germline-specific *nanos* (*nos*) promoter, *vasa* promoter, or the ubiquitous *Actin5c* (*Act*) promoter (Kondo and Ueda, 2013; Ren *et al.,* 2013; reviewed in: Port *et al.,* 2014).

Though CRISPR/Cas9 is regarded as a good system for its target specificity (reviewed in: Wu *et al*., 2014), there are limitations when using the system. The requirement for a PAM motif sequence to be immediately next to the 3' end (Jinek *et al.,* 2012) dictates what sequences can be targeted. In addition, there is the potential for species with lots of PAM sites in their genome to have more off-target alterations, making off-target effects possible in *D. melanogaster* and *D. simulans* (Wu *et al*., 2014). It is still difficult to predict genome-wide off-target effects of Cas9. For example, a recent study done in mice found that sgRNAs target loci independently of their target, causing a high number of off-target CRISPR-induced indel mutations that could be deleterious (Schaefer *et al*., 2017). It is still being worked out whether improving sgRNA design or using a high-fidelity Cas9 could reduce off-targets, meaning that the general rules for optimizing Cas9/sgRNA specificity are still incomplete (reviewed in: Wu *et al*., 2014).
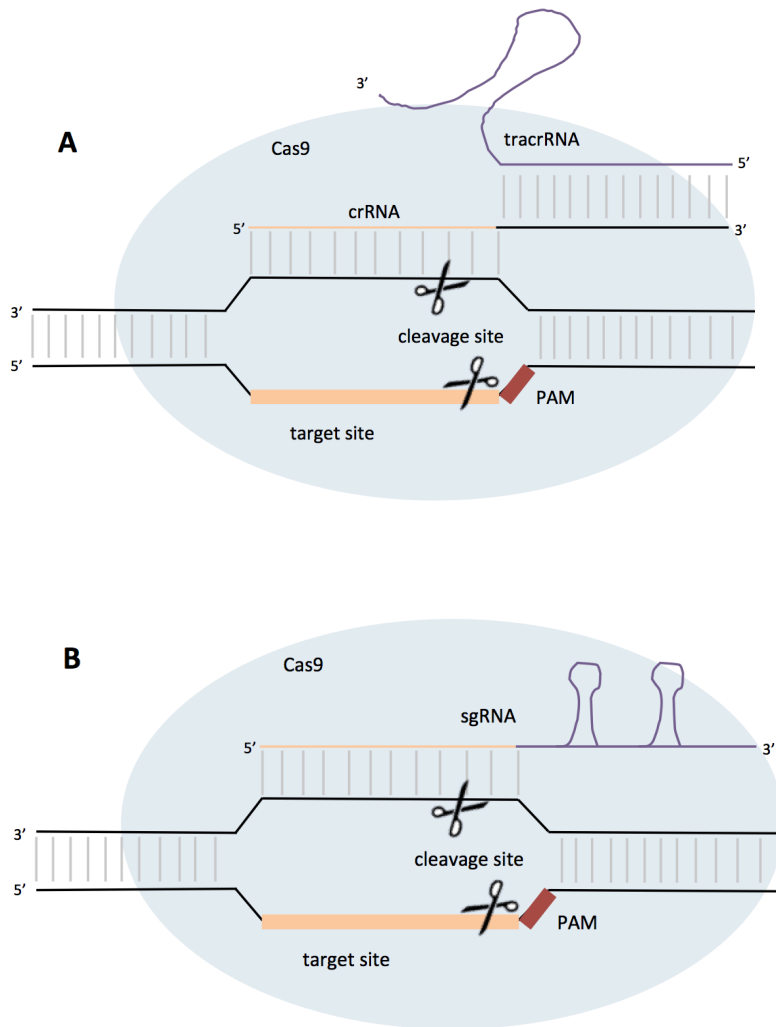
Figure 6: CRISPR/Cas9 type II system.

A. Schematic of type II CRISPR/Cas9 system in bacteria. Targeted (thick light orange line) cleavage (scissors) of viral DNA adjacent to a PAM motif (red) is achieved through the formation of a crRNA (thin orange line) and tracrRNA (thin purple line) complex with Cas9 endonuclease (blue oval). **B.** Schematic of two-component type II CRISPR/Cas9 system for use *in vivo*. The Cas9-sgRNA complex shows how the sgRNA (thin purple and orange line) is designed to target the complimentary sequence of the target site (thick orange line), adjacent to the PAM sequence (red), allowing for the Cas9 endonuclease (blue oval) to cause a double strand break at the cleavage site (scissors). This figure was adapted from Bassett and Liu (2014).

1.5.3   Repairing DSBs

Cleavage of a target sequence triggers repair mechanisms within the cell to repair a DSB. This can take place in two ways, either though non-homologous end joining (NHEJ) or homologous recombination (HR) (Figure 7). NHEJ is a common repair pathway in multicellular organisms that is used to repair stress-induced DSBs as well as DSBs that were generated with enzymes, such as when using the CRISPR/Cas9 system. The chromosome is repaired without the use of external homologies, instead relying on the modification of the broken ends to make them compatible in order to bring the ends together to re-join the sequence using ligase IV (in *Drosophila*) (Lieber *et al*, 2003; McVey *et al*., 2004). This often results in deletions of 1-10 nucleotides from each end of the DSB, resulting in indel mutations that can cause a frameshift in the sequence (Figure 7A; Lieber *et al*., 2003). If two DSBs are made within the same sequence, repair through NHEJ can result in a deletion of the sequence between those two break points (Figure 7B).

HR, on the other hand, uses a homologous chromosome or homologous donor template to copy information from one homologue to the other (Haber, 1995; Lieber *et al*., 2003). Where this process naturally occurs in the yeast *S. cerevisiae*, the repair of DSBs occurs without the loss of genetic information (Haber, 1995). HR is also being widely used in transgenics to either insert a gene (Figure 7C) or replace a donor template with an alternative sequence (Figure 7D) at the target where the DSB was generated (reviewed in: Bassett and Liu, 2013; reviewed in: Housden *et al.,* 2014; reviewed in: Ren *et al*, 2014). Both repair mechanisms are native to *Drosophila* and so either option is available when choosing a means to repair a DSB generated by CRISPR (reviewed in: Bassett and Liu, 2013; reviewed in: Housden *et al.,* 2014; reviewed in: Ren *et al.,* 2014). The repair mechanism of choice depends on the preference for either the creation of small deletions or precise genome modifications. I will be generating two DSBs in exons of *fru* using CRISPR/Cas9-mediated NHEJ to cause a deletion in the sequence.
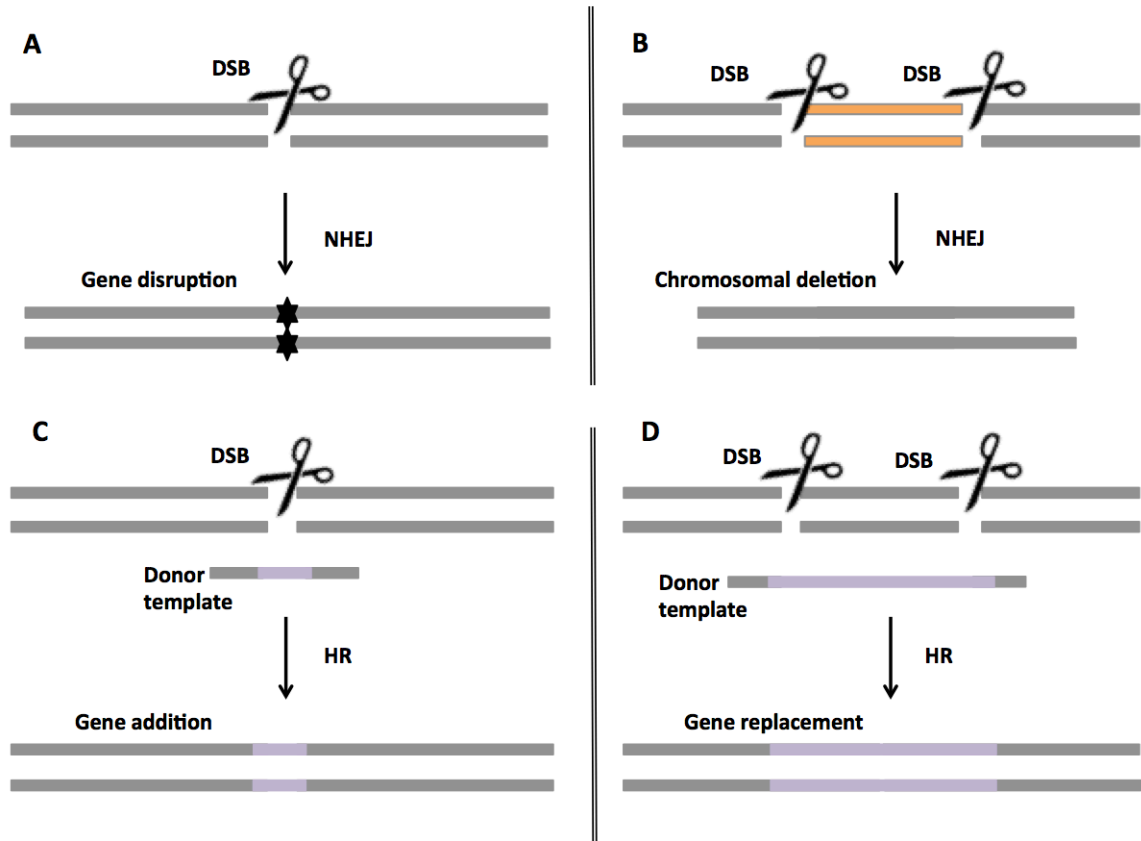
Figure 7: Representation of repairing DSBs with (A, B) NHEJ and (C, D) HR.

A When a single DSB is made in a sequence, repair with NHEJ can result in a gene disruption due to the possibility of a few nucleotides during the repair. **B** If two DSBs are made in the same sequence, repair through NHEJ can result in a chromosomal deletion of the portion of the sequence in between the two break points. **C** For repair to a single DSB by HR, a donor template with flanking sequence that is homologous to the sequence on either end of the cut can be added to the sequence during repair, resulting in the addition of a gene. **D** The occurrence of two DSBs in a sequence allows for the possibility of replacing one gene with another by use of a donor template paired with a sequence homologous to the donor sequences flanking the cut sites.

1.5.4   Overview

My goal is to confirm definitively that *fru* has a molecular role in female rejection behaviour by generating reciprocal gene knockouts in *D. melanogaster* and *D. simulans* using CRISPR/Cas9-mediated NHEJ, targeted to the common regions in *fru*. Since I do not know which *fru* transcript underlies female rejection behaviour, I can ensure that *fru* expression is completely disrupted by targeting the common region, which is present in all *fru* transcripts. I aim to generate DSBs in both exons C1 and C2, or in C4 and C5, allowing for the intervening portion of sequence to be removed. Removing two separate regions reduces the likelihood that any observed phenotypes are due to off-target effects.

Knocking out *fru* will allow for the generation of *D. melanogaster*/*D. simulans* hybrids with either the *D. melanogaster* or the *D. simulans fru* allele disrupted. These hybrids will be genetically identical throughout the genome except at the *fru* locus, where only the allele of a single species will be expressed. This will allow for the performance of a reciprocal hemizygosity test (Figure 5; Stern, 2014) that can positively identify if *fru* contributes to the differences in female behaviour between the two species. In this experiment, two types of *D. melanogaster*/*D. simulans* hybrids will be generated and tested to confirm the effect of *fru* on species-specific female rejection behaviour. In hybrids that only express the *D. simulans* allele of *fru*, females are expected to reject *D. melanogaster* males and have reduced mating, while in the reciprocal cross, a female that only expresses the *D. melanogaster fru* allele, will not reject *D. melanogaster* males (Figure 8). This experimental approach will verify that a gene contributes to the genetic basis of species-specific female rejection behaviour. From there, it may be possible to map out the genetic interactions of the gene with other loci that affect female mate preference and then locate the neurons in which these genes are expressed, forming signals to elicit female preference behaviours. Altogether this will tell us how behavioural isolation acts at the molecular level to give rise to a phenotypic barrier of speciation.
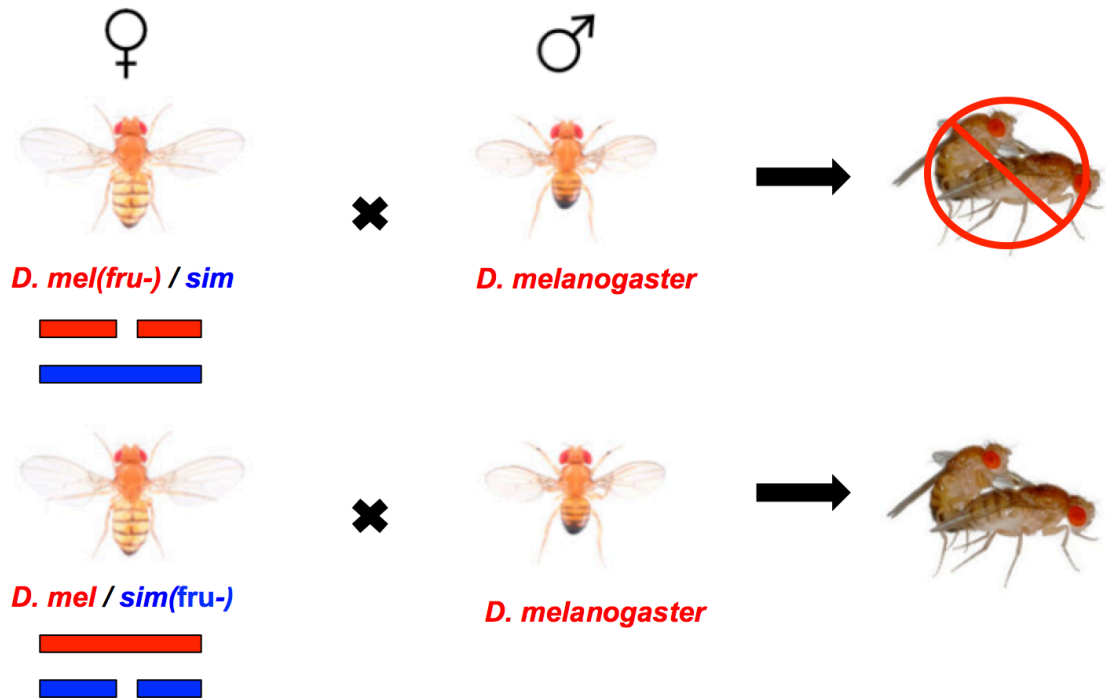
Figure 8: Reciprocal hemizygosity test between _D. melanogaster_ and _D. simulans_ with _fru_ disruption.

Female hybrids containing one homolog from _D. melanogaster_ (red) and _D. simulans_ (blue) that have a species-specific allele of _fru_ disrupted (broken lines). These females are crossed with _D. melanogaster_ males to see if the female will accept or reject mating. As _fru_ has been previously shown to affect female preference, the behaviours exhibited by these two hybrids are expected to be different, confirming that _fru_ has a role in female species-specific rejection behaviour.

2 Materials and Methods

2.1 Fly husbandry

Drosophila stocks were maintained on standard cornmeal medium (Bloomington *Drosophila* Stock Center recipe) in 30 mL vials at 24°C, 70% humidity with a 14:10 light:dark cycle. Wild-type *D. melanogaster* line BJS was obtained from Dr. Brent Sinclair. A transgenic *D. melanogaster* line expressing *Cas9* under the *nanos* promoter on the X chromosome (y^1 *M{nos-Cas9.P}ZH-2A* w^*) and a balancer stock (w^{1118}/*Dp(1;Y)y$^+$; CyO/nub^1 b^1 sna^{Sco} lt^1 stw^3; MKRS/TM6B, Tb^1*) were obtained from the Bloomington *Drosophila* Stock Center. Wild-type *D. simulans* line Florida City (FC) was obtained from Dr. Jerry Coyne. A *D. simulans* stock with an inversion on the third chromosome (*D. simulans Dlp/st Ubx*) and a transgenic *D. simulans* stock expressing *Cas9* under the *nanos* promoter on the *X* chromosome (*D. simulans 1029::Cas9*) were obtained from Dr. David Stern. A *D. melanogaster* GFP-tagged sperm line was obtained from Dr. John Belote.

2.2 CRISPR constructs and sgRNA transcription

Target sites consist of 20 nucleotides with an additional three nucleotide (NGG) protospacer adjacent motif (PAM) sequence at the 3' end, essential for Cas9 binding and cleavage of the target sequence. Optimal target sites have one or two guanines (G) at the 5' end of the target sequence with no or minimal off-target sites. Target sites for sgRNA chosen for common region exons in *fru* (C1, C2, C4, and C5) for *D. melanogaster* and *D. simulans* (Appendix A) were identified using the flyCRISPR optimal target finder (Gratz, *et al.,* 2014; http://tools.flycrispr.molbio.wisc.edu/targetFinder/). As a control for the effectiveness of my sgRNA generation protocol and injection protocol, I used sgRNA for *frost* (*fst*) and *yellow* (*y*). *fst* sgRNA was *in vitro* transcribed by Alaa Briek (courtesy of the lab of Dr. Anthony Percival-Smith) and maintained at -80°C, and sgRNA target site for *y* was taken from Bassett and Liu (2014) and *in vitro* transcribed by myself.

Unlike the *D. melanogaster fru* sequence, *D. simulans fru* sequence was not yet annotated to identify the different exons of *fru*, therefore requiring an extra step before target sites

could be determined. To find *fru* targets in *D. simulans*, I annotated the *fru* sequence in comparison to the *D. melanogaster fru* sequence using BLAST and Clustal Sequence Alignment (EMBL-EBI). I amplified the common region exons in *D. simulans* FC samples through PCR using exon specific primers, followed by gel extraction for clean-up using the Geneaid Purification kit (FroggaBio, North York, ON). Samples were sent to the Robarts DNA Sequencing Facility (London, ON) for sequencing to ensure target sequences were present in *D. simulans* FC. Though the common region is mostly conserved between the two *Drosophila* species, SNPs were present, and thus the sequence of the particular strain was useful in identifying precise common region targets for sgRNA for both *D. melanogaster* and *D. simulans* (Appendix A; Table 1).

sgRNA for the four *fru* targets in *D. melanogaster*, 10 *fru* targets in *D. simulans,* and the one *y* target, were generated through the *in vitro* transcription method outlined by Bassett and Liu (2014). Firstly, target sequences had to be amplified to use for *in vitro* transcription of sgRNA. This was done by amplifying CRISPR F oligonucleotides containing the target sequence (without the PAM sequence), T7 promoter and gRNA backbone (Tables 1) alongside the CRISPR R oligonucleotide (AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTAT TTTAACTTGCTATTTCTAGCTCTAAAAC) (Eurofins, Mississauga, ON). When designing CRISPR F oligonucleotides, if the target sequence did not begin with guanine, then a guanine could be added to the 3' end of the T7 promoter sequence (Table 1; Bassett and Liu, 2014).

After amplification of the CRISPR oligonucleotides, 2 μL of PCR product sgRNA template was analyzed on a 2% TBE agarose gel to ensure a single band. The remaining product was purified using a Geneaid Purification kit (FroggaBio, North York, ON) and 300 ng was used for *in vitro* transcription. After initial *in vitro* transcription of sgRNA using a T7 MEGAscript kit (ThermoFisher Scientific, Burlington, ON), sodium acetate was added to 10.5 ng of sgRNA, followed by 70% ethanol to precipitate the solution before analyzing the product on a 1.2% TAE agarose gel and freezing at -80°C until needed for the injections.

Table 1: Target sites and sgRNA oligonucleotides for *fru* exons C1, C2, C4, and C5, and *y*

Species	Target genes and exon	CRISPR F Oligonucleotides (5' to 3')		
		T7 Promoter	Target site	gRNA backbone
D. melanogaster	*fru* **C1**	GAAATTAATACGACTCACTATAGG	CAACACTGACCAAGGAGCGA	
	fru **C2**	GAAATTAATACGACTCACTATAGG	TGCAGGCTGACAGGATGGTC	
	fru **C4**	GAAATTAATACGACTCACTATAG	GCTCAGACCGCTAAGGGCGC	
	fru **C5**	GAAATTAATACGACTCACTATAGG	CTCGAACTGATACAGCCAGT	
	y	GAAATTAATACGACTCACTATAG	GTTTTGGACACTGGAACCG	
D. simulans	*fru* **C1**	GAAATTAATACGACTCACTATAGG	CAACACTGACCAAGGAGCGA	GTTTTAGAGCTAGAAATAGC
	fru **C2**	GAAATTAATACGACTCACTATAG	GCTGCCCATGTTTCTCAAGA	
	fru **C4a**	GAAATTAATACGACTCACTATAGG	TCGAGTCCCGTGCCCAAAAC	
	fru **C4b**	GAAATTAATACGACTCACTATAG	GGCGCGGCCAGCGCCCTGAG	
	fru **C4c**	GAAATTAATACGACTCACTATAGG	TCGGAGTCGGAGGACGCCGG	
	fru **C4d**	GAAATTAATACGACTCACTATAG	GCCGCCCAGATGGACGCTGG	
	fru **C5a**	GAAATTAATACGACTCACTATAGG	CTCGAACTGATACAGCCAGT	
	fru **C5b**	GAAATTAATACGACTCACTATAG	GCAGCAGCCACACCTTACGC	
	fru **C5c**	GAAATTAATACGACTCACTATAGG	CAAAGCCCAGCGATACCATC	
	fru **C5d**	GAAATTAATACGACTCACTATAGG	TCGCAAATGGAGCGGTGAGC	

2.3   Egg collection

Between one week to two days before injection, flies were placed into fly cages with a grape or apple juice agar plate (Recipe in Appendix B) smeared with yeast; plates were changed twice a day when injections were not being performed. The day before injecting, plates were changed every hour during the window of time injections would be performed the next day (e.g. if injections were to be done between 12 p.m. and 6 p.m., plates would be changed every hour between 12 p.m. and 6 p.m. the day before). This was done to help induce flies to lay eggs on the juice plates during the time injections would be performed. On the day of injections, juice plates containing eggs were collected every 30 minutes and replaced with a fresh juice plate. When preparing eggs for injection, eggs could either be left with their chorion on, or have their chorion removed.

2.3.1   Egg collection with bleach dechorionation

Dechorionation of eggs was done by soaking the eggs on the juice agar plate in 3% bleach for 1 minute, pouring the bleach with the eggs into a mesh egg basket and rinsing with deionized water. Dechorionated eggs were then transferred from the mesh onto a fresh juice agar plate (no yeast) to allow for easy manipulation of the eggs. All eggs were vertically aligned with posterior ends in one direction with a fine paintbrush. Aligned eggs were then transferred onto a cover slip using double-sided sticky tape and then mounted on a microscope slide. Eggs were desiccated by placing under a hairdryer for an appropriate time (in a 18°C room, embryos were dried for 4.5 minutes; at 24°C, 3 minutes), allowing for better survivability and higher likelihood of sgRNA entering the egg while decreasing leakage of cytoplasm during injection. Eggs were then covered in halocarbon oil. Initial tests to find the most efficient dechorionation protocol were done in partial assistance with Melissa Wong who contributed to half of the results for testing survivability after wire-and-tape dechorionation and injections (see Results).

2.3.2   Egg collection without dechorionation

Chorionated eggs were washed off a juice agar plate with water and poured into a mesh basket. They were then transferred to a water droplet on a coverslip mounted on a

microscope slide and vertically aligned with posterior ends in one direction, but at a slight angle, using a fine paintbrush. Eggs were given time to dry so that they could adhere to the coverslip; adherence was determined by pushing the eggs lightly with a dry paintbrush to see if they would move. Eggs were then covered in halocarbon oil.

2.4   Microinjection and screening mutations

Before and after injections, sgRNA was electrophoresed on 1.2% TAE agarose gel to ensure that the sgRNA had not degraded. Injections were performed using a glass capillary tube injection needle made with the needle puller (Sutter Instrument Company, Micropipette Puller P-97) provided by Dr. Gregory Gloor, using protocol 9 on the machine. Injection needles were loaded with sgRNA in an injection mix, sgRNA alone, or sgRNA and Cas9 using glass capillary loading needles on the day injections were being performed. The needles were then mounted into a needle holder on a dissecting microscope (Nikon Stereo microscope) attached to a digital microinjector (Sutter Instrument Xenoworks Digital Microinjector). Eggs were injected at the posterior end and drained of halocarbon oil after injections. The injected dechorionated eggs were then transferred to a food plate smeared with yeast, kept at 18°C for 48 hours, and transferred to a 24°C incubator with a 14:10 light:dark cycle. Eggs injected without dechorionation were put into a beaker with wet kimwipes and sealed with parafilm (Figure 9). They were then placed in a 24°C incubator with a 14:10 light:dark cycle and removed from the beaker after 48 hours. Once the dechorionated or non-dechorionated eggs developed into larvae, the larvae were transferred to 30 mL food vials where they could eclose into adults.
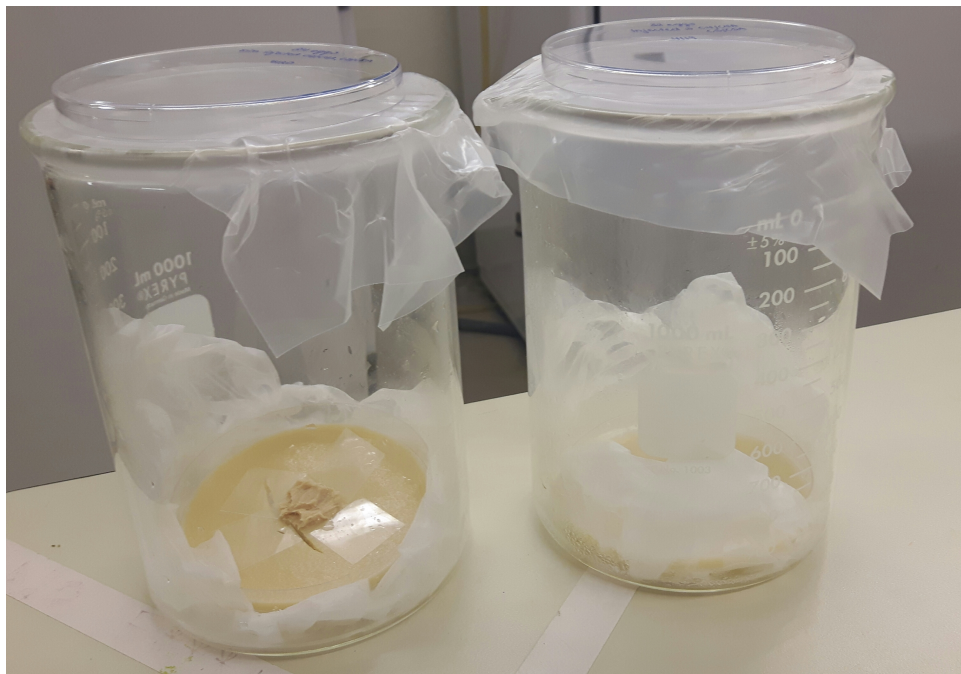
Figure 9: Incubation of injected eggs with chorion in beakers sealed with parafilm.

After injection, the coverslip containing the injected eggs is placed on a food plate smeared with yeast (seen here in the beaker surrounded by kimwipes) angled slightly towards the yeast in the middle to allow for extra halocarbon oil to drain and for developing larvae to gain easier access to the food. Wet kimwipes surround the food plate to provide moisture to the developing embryos. After 48 hours, food plates are removed from the beakers.

2.4.1   Injections for targeted *fru* mutations

Eggs were injected with either: sgRNA in an injection mix (250ng/μL of each sgRNA, 500ng/μL repair template, 0.5 μL food dye, 2 μL 10% glycerol, and 2 μL 1X PBS), or sgRNA alone, if injecting into *D. melanogaster* eggs expressing *Cas9*, or sgRNA and Cas9 protein (New England Biolabs, Ipswich, Massachusetts) if injecting into *D. simulans* FC. When injecting, two sgRNAs targeting two common region exons of *fru* were paired such that there could be the creation of a DSB in both C1 and C2, or in C4 and C5. Best Gene and I performed injections of sgRNA targeting C1, C2, C4, and C5 in *D. melanogaster* in duplicate.

2.4.2   Crosses and genotyping for *fru* mutations

Injected flies with targeted *fru* disruption (G$_0$) were mated to either a balancer (*D. melanogaster*) or an inversion (*D. simulans*) stock to allow for maintenance of a CRISPR-modified *fru* locus. After mating, injected adults (G$_0$) were scored for successful gene disruptions through genotyping using PCR with primers flanking the cut sites, followed by restriction enzyme (RE) digestions (Table 2; Figure 10). To ensure that the deletion was in the germline, the offspring of parents with mutations (G$_1$ – heterozygotes for *fru* deletion over a balancer or inversion) were then crossed again to balancer or inversion stock flies and G$_1$ flies were genotyped and scored again for the gene disruption. The offspring (G$_2$) of G$_1$ flies bearing a disruption were crossed together, allowing for the generation of a stable stock of flies with successful *fru* disruptions maintained over a balancer or inversion, as a homozygous *fru* disruption is lethal (Ryner *et al.*, 1996). Flies injected by Best Gene were crossed to *D. melanogaster* Cas9 stock and balancer stock, making the third chromosome genotype *fru*$^{C4-}$*/TM6B, Tb*^{1}*,* with the X chromosome recombinant between that of the Cas9 stock and the balancer stock, and the second chromosome recombinant between that of the Cas9 stock and the *Scutoid*-bearing homolog of the balancer stock.

Screening for knockouts was done either by amplification with primers flanking two exons and their cut sites, to screen for a large deletion of these exons, or with primers specific to each exon such that the portion of the sequence surrounding the Cas9 cut site

could be amplified and used for RE digest, to identify a deletion in a single exon (Table 2). Samples were amplified under the following PCR conditions when using the Taq polymerase from Invitrogen (ThermoFisher): 1 cycle 95°C 3 min; 2 cycles 95°C 30 s/58°C 30s/72°C 30 s; 4 cycles 95°C/30 s 57.2°C/30s 72°C/30 s; 30 cycles 95°C/30 s 56.5°C/30s 72°C/30 s; 1 cycle 72°C/5 min; hold 4°C. When using Froggabio Taq polymerase extension time was longer, as follows: 95°C 3 min, 95°C 30 s, 58°C 30s, 72°C 1 min (2 cycles), 95°C 30 s, 57.2°C 30s, 72°C 1 min (4 cycles), 95°C 30 s, 56.5°C 30s, 72°C 1 min (30 cycles), 72°C 10 min, 4°C hold. Samples were then electrophoresed on 2% TBE agarose gels to screen for deletions according to product size.

When screening for deletions in only a single exon, REs were used that had a recognition sequence that overlapped the Cas9 cut site. These REs were identified using NEBcutter (Vincze *et al.*, 2003; http://www.labtools.us/nebcutter-v2-0/) (Table 2). RE digests were prepared using 10 µL of PCR product, 18 µL of nuclease-free water, 2 µL of appropriate buffer (provided with RE), and 10 units of RE (generally 0.5 µL). This mix was incubated at the appropriate temperature for the specific enzyme for 13-16 hours, denatured for 20 minutes, and held at 4°C until samples could be electrophoresed on a 2% TBE agarose gel to visualize bands (Figure 10). When mutations were detected, samples were gel extracted, purified using the Geneaid Purification kit (FroggaBio), and sent to the Robarts DNA Sequencing Facility (London, ON) for sequencing to identify the extent of deletions. Amino acid sequences for the common region were obtained from FlyBase (Gramates *et al.*, 2017) and NCBI BLAST (Altschul, *et al.*, 1990) to identify amino acid changes due to deletions and variation in sequences amongst *Drosophila* species. Protein structures were modeled using Phyre2 (Kelley, *et* al., 2015; http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) default parameters. The default parameters of Phyre Investigator and SuSPect were ran to assess alignment quality and mutational analysis.

Table 2: Primers used for genotyping by PCR for single or double cut-site induced mutations in *fru* along with RE.

	Target exon	Primer sequence (5' – 3') Forward and Reverse	Restriction enzyme
D. melanogaster	C1-C2	**F:** TGTTTCGCAAGACTCGCTT **R:** ACACCGTTACACACGACCAA	
	C4-C5	**F:** TGTTGGGTCCTCTTTCAAGG **R:** GCCAACTTCAAGTCGAGTCC	
	C1	**F:** TATCTCATGGACGCACCTTG **R:** AAACAAAGCAGGCGCTAAAC	*BseDI*
	C2	**F:** GATCCCTGATTTGCACACAC **R:** CAAGGTGCGTCCATGAGATA	*DrdI*
	C4	**F:** ACGATGCAGCAACAAAATCA **R:** TCCCTTTCAATGGCAGACTC	*Bpu10I*
	C5	**F:** GCTAATCCTGAGCGGTGTGT **R:** GAAGCGTTTTTAGGCAGCAC	*Bsr1*
D. simulans	C1-C2	**F:** ATGCTCTTGTCTCGCCACAT **R:** AGTCGGAGCGGTAGTTCAGA	
	C4-C5	**F:** GGTTTCGCGTCGTTATCAGT **R:** GCTAATCCTGAGCGGTGTGT	
	C4	**F:** GGTTTCGCGTCGTTATCAGT **R:** CTCACGTTTTTGGGAAGCAT	*Bpu10I* *NaeI*
	C5	**F:** GAAGCGTTTTTAGGCAGCAC **R:** GCTAATCCTGAGCGGTGTGT	*Bsr1* *LpnP1*
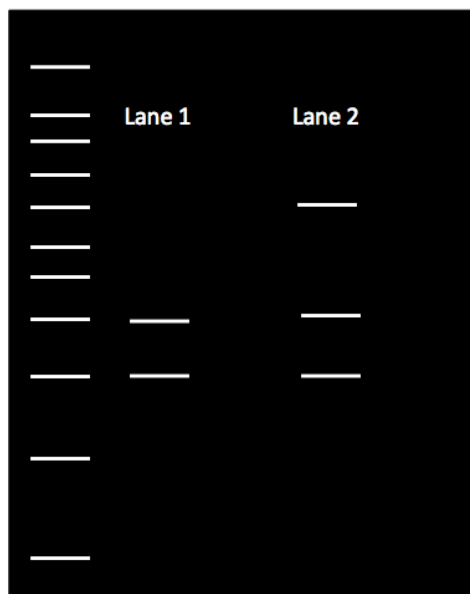
Figure 10: Illustration of genotyping for mutations using a RE digestion.

When performing a RE digestion to determine if a single sgRNA was able to target a specific exon in *fru*, the RE would only fully digest if the sequence was not targeted, indicating Cas9 did not make a DSB. Lane 1 shows what that digest would look like – the full sequence has been digested, represented by the two shorter fragments. Lane 2 shows the digest if sgRNA was capable of targeting the sequence, causing a DSB to occur. Partial digestion would still occur (as seen with two smaller fragments), as CRISPR would not be able to cause a DSB in every cell. The larger fragment on top represents the portion of the sequence that was not digested due to a mutation caused by the DSB through Cas9. G_0 flies with CRISPR-mediated mutation would look like Lane 2.

2.4.3 Injections and screening for *fst* mutations

Control injections using two sgRNAs targeting *fst* and a repair template with a *white*[+] marker (courtesy of the lab of Dr. Anthony Percival-Smith) were mixed in an injection solution of: 250ng/μL of each sgRNA, 500ng/μL repair template, 0.5 μL food dye, 2 μL 10% glycerol, and 2 μL 1X PBS. Food dye was used to help visualize the solution when it was being injected. Transgenic *D. melanogaster* expressing Cas9 under the *nanos* promoter were used for *fst* injections. Crossing viable injected adults to wild-type *D. melanogaster* and screening progeny for white eyes was done to screen for *fst* mutants.

2.4.4 Injections and screening *y* mutations

Injecting sgRNA targeting *y* used no injection mix as described above, but only sgRNA. As well, there were different flies than those used for targeting *fru* common regions in *D. melanogaster*. The original stock used for *fru* injections contained a mutation in the *y* gene, which would prevent scoring for additional CRISPR-induced mutations in this gene. Therefore, hybrid transgenic *Cas9 D. melanogaster/*wild type *D. melanogaster* males or females were used, as these hybrids would contain a single functional allele of *y* on the X chromosome, which could be targeted for disruption. To generate hybrids for injections, virgin transgenic Cas9 females and wild-type males (or vice versa) were put into fly cages such that the eggs laid were hybrids. Once grown, hybrid female injected flies (G_0) were selected for and mated to wild-type male *D. melanogaster*. Their male progeny (G_1) were then visually screened for *y* disruption (Figure 11). If the injections worked and *y* was disrupted, all males would lack *y* expression, as seen by yellow discoloration of their bodies.
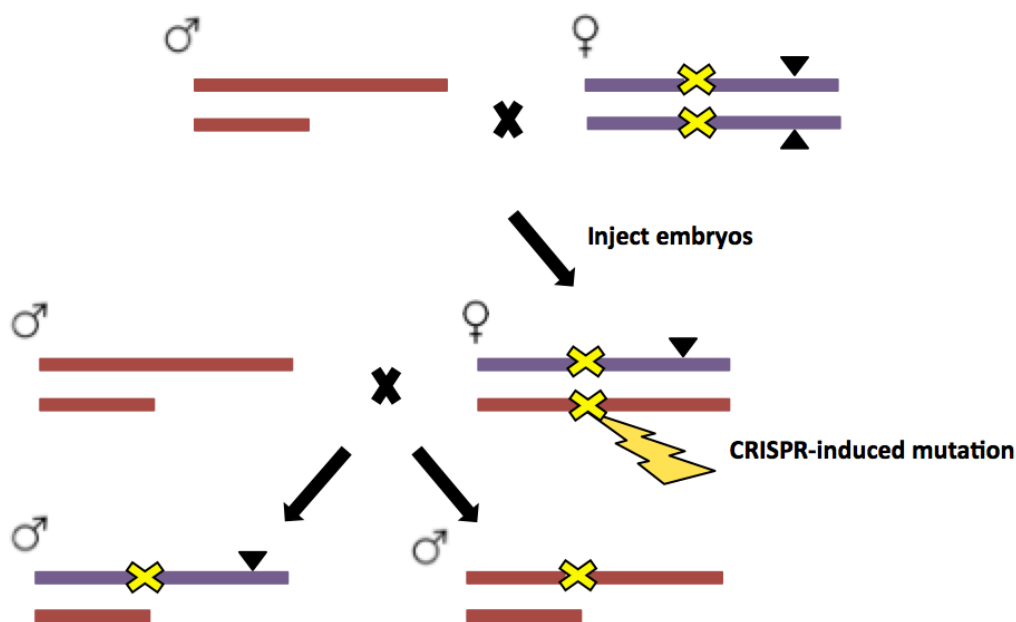
Figure 11: Crossing scheme for inducing *y* mutation in transgenic hybrids and scoring for *y* disruption.

The bars in this figure represent chromosome 1 (X and Y) in *Drosophila*, where two bars represent one individual fly with a set of chromosomes; two equal length chromosomes are females (two X chromosomes); one long (X) and one short (Y) chromosome are males. Yellow x's represent a *y* disruption. Red bars represent wild-type homologs in wild-type *D. melanogaster* flies; purple bars are homologs containing the *Cas9* transgene (black triangle) and *y* disruption (yellow x) in transgenic *D. melanogaster*. In the first cross, a male wild-type *D. melanogaster* is mated to a female transgenic *Cas9 D. melanogaster*. Hybrid eggs produced by this cross are used for injections and ideally would produce a female with the inherited copy of *y* disruption from the mother and a disrupted *y* copy as generated by CRISPR (lightning bolt). To ensure *y* was disrupted by CRISPR, this female is mated to a wild-type male (second cross), and her male progeny are screened for *y* disruption. If all males produced from a single injected female exhibit *y* phenotype, and not just a portion of males (as they merely inherited the disrupted copy from the transgenic allele), then CRISPR worked in the *y* control injections (male progeny shown).

2.5   Behavioural mating assays

D. melanogaster with CRISPR-mediated NHEJ deletions in C4 (see Results) were balanced over *TM6B* to make three stable stocks. The first stock to be stable had a three-amino acid deletion (deletion stock C4-a) (see Results) and was used to test male and female mating behaviour.

2.5.1   Male behavioural sterility assay

Males from *fru* deletion stock C4-a were used to test male sterility, as males homozygous for *fru* mutations do not attempt to copulate with females (Hall, 1978). The protocol used to test for male behavioural sterility was taken from Gailey and Hall (1989). Briefly, pupae that were homozygous for the *fru* C4 deletion were collected; virgin males were collected at eclosion and kept with no more than 10 males in one vial. Wild-type virgin males from the stock used for injections were also collected to act as a control for male behaviour. All males were aged 5-7 days before placing individually into a food vial with up to 4 virgin wild-type females from the injection stock, aged 5-7 days. After 7 days, vials were scored for presence of larval progeny.

2.5.2   Female receptivity behavioural mating assay

D. melanogaster with deletion in C4 exon of *fru* was also tested to see if the deletion had an effect on female species-specific mating preference. Virgin females heterozygous for C4 deletion maintained over *TM6B* balancer were collected at eclosion and aged for 5-7 days before crossing to wild-type virgin *D. melanogaster* males from the Cas9 stock used for injections or wild-type virgin *D. simulans* FC males, aged 5-7 days. From each cross, virgin females with the balancer and virgin females with the deletion were collected and aged 5-7 days. The four different genotypes of females were then used for behavioural mating assays to see if the deletion of C4 in hybrid females would cause a reduction in mating compared to the *D. melanogaster*/*D. simulans* hybrid with the balancer and the *D. melanogaster* females heterozygous for the deletion or heterozygous for the balancer (as in Figure 2). Behavioural mating assays were performed by observing courtship and mating between virgin female and virgin wild-type *D. melanogaster* males with GFP-

tagged sperm in 30 mL glass vials for one hour at 24°C with 70% relative humidity. After one hour, the pairs were placed into 30 mL food vials for 24 hours. After 24 hours, the female reproductive tract was dissected to score for presence of sperm (as a proxy for mating) or absence of sperm (as a proxy for no mating) to determine if there was a change in behavioural phenotype to reject a mate. The proportion of pairs that mated was analyzed using a G-test. If a reduction in mating of the hybrid females with the deletion in C4 exon were observed, than this line would be appropriate to use for one half of the reciprocal hemizygosity test (Appendix C).

3   Results

The goal of this study was to knockout *fru* in *D. melanogaster* and *D. simulans* in order to confirm the role of *fru* in female species-specific mate preference behaviour. To achieve this, first reciprocal gene knockouts in *D. melanogaster* and *D. simulans* had to be achieved by targeting the common region exons in *fru* to ensure gene disruption using CRISPR/Cas9 and NHEJ repair of the sequence. This approach was not entirely successful, as deletions in only a single exon of *fru* were achieved in *D. melanogaster*. Therefore, this study did not see the completion of generating female hybrids expressing only one species-specific allele of *fru* and testing for mating behaviour with a *D. melanogaster* male using the reciprocal hemizgosity test.

3.1   Testing injection protocols: survivability of eggs

A key component to having CRISPR/Cas9 target the gene of interest is to be able to inject the CRISPR/Cas9 components into appropriately aged eggs. An injection protocol had not yet been set up in the Moehring lab at the start of this project, but through learning different injection protocols and adapting them for the lab space and equipment available, I was able to establish an injection protocol for the Moehring lab.

An injection protocol can be broken down into the following components: preparation of injection and loading needles, preparation of flies for egg laying, dechorionation (removal of the chorion) of eggs, desiccation of eggs, and microinjection of eggs. To establish an efficient injection protocol, troubleshooting of dechorionation and microinjection were first completed. Removal of the chorion allows for easier puncturing of the embryo with a needle, but the process of removal can cause reduced survivability. Though some injections were done with dechorionation, injections with an intact chorion were a better approach for this study (see below). Other approaches that were tested were the use of an injection mix (see Results section detailing *y* injections) and different strategies for egg laying (see Discussion).

The purpose of dechorionation is to remove the outer shell from the egg and allow access to the embryo, making it easier to identify the age of the embryo and easier for the

injection needle to pierce the embryo. The two different dechorionation methods tried were: washing eggs in 3% bleach for one minute as this was sufficient time to remove the entire chorion (personal communication, Dr. Anthony Percival-Smith) and lining up eggs between two thin wires and removing the chorion with sticky tape (personal communication, Dr. Anne Simon). Another option that was later explored was to leave the chorion intact (Bassett and Liu, 2014). These tests were carried out on eggs collected from either transgenic *D. simulans Cas9*-expressing flies or transgenic *D. melanogaster Cas9*-expressing flies.

Of the two dechorionation methods tested on *D. simulans Cas9*-expressing eggs, 3% bleach was deemed to be the most effective (when first attempted at the beginning of this project), since bleach preparation was faster, even though it had a significantly lower survivability of eggs than the wire and sticky tape method (Table 3; 49% survivability with bleach vs. 100% survivability with tape; z-test, P = 0.001).

After initial injections of dechorionated eggs with saline solution, bleach also had a lower survivability (z-test, P = 0.0459) than the wire-and-tape method, but the problem with using wire and sticky tape was the eggs would often be shifted out of orientation. This presented a problem since all eggs must be in a straight line with their posterior ends facing one direction in order for microinjections to be quick and effective. Realigning the eggs was difficult and time-consuming, surpassing the 10-minute interval allotted to dechorionation and alignment (Table 4). Therefore, the strategy of dechorionating eggs with bleach was chosen for injections targeting disruptions in *fst, y,* and *fru* in *D. melanogaster*.

Later in the year, injections of *D. melanogaster Cas9*-expressing eggs after bleach dechorionation had lower survivability (Table 7, 8), and so bleach dechorionation was retested on *D. melanogaster Cas9*-expressing eggs and found to have significantly lower survivability than when first tested (z-test, P = 0.00002). The dramatic shift in bleach survivability between previous tests (Table 3) and more current survivability tests (Table 5) make dechorionation an inconsistent approach. A factor like change in potency of bleach over time could have played a role, but was not tested. I found an alternative

approach for preparing eggs for microinjections, which was to keep the chorion on and skip the dechorionation process. This approach was tested to see if there would be better survivability. When egg survivability was tested with bleach treatment and without the dechorionation process (without injection) on transgenic *D. melanogaster* flies, it was found that keeping the chorion on was a better approach to ensure significantly higher survivability (z-test, P = 0.000001) before injections (Table 5). However, survival of eggs with an intact chorion was expected to be 100% as there was no dechorionation or injection of the eggs, but there was only 50% survivability. These tests used *D. melanogaster* Cas9 line, as opposed to transgenic *D. simulans* Cas9, which was used in the original dechorionation survivability tests. Low survivability could indicate there is a low eclosion rate amongst the transgenic *D. melanogaster Cas9*-expressing flies compared to the *D. simulans Cas9*-expressing flies.

Additional steps in the microinjection protocol including desiccating the eggs after dechorionation, or covering the eggs with halocarbon oil, were also tested for survivability (Table 6). These tests were all seen to lower survival, although the survivability difference between dechorionating with bleach alone, with addition of halocarbon oil (z-test, P = 0.737), or desiccation time (z-test, P = 0.626) was not significant. Survivability was significantly lower for intact eggs covered in halocarbon oil compared to those that were not covered with halocarbon oil (z-test, P = 0.023), but survivability of eggs with chorion on and halocarbon oil was significantly higher than eggs exposed to bleach and halocarbon oil (z-test, P = 0.029). Therefore, skipping the dechorionation process was seen as a way to ensure more consistent survivability and potentially higher survivability when injecting. When performing injections to target *y*, survivability after injections using bleach dechorionation and no dechorionation of *D. melanogaster* eggs were compared (see section 3.3.1) and survivability was not significantly different for the two approaches (z-test, P = 0.590). Injecting with an intact chorion was less time consuming than dechorionation, so the chorion intact protocol was employed for injecting into *D. melanogaster* and *D. simulans* to disrupt *fru.* However, this approach could have limited mutation efficiency, as it was more difficult to visualize where sgRNA was being injected into the egg.

Table 3: Survivability of transgenic *D. simulans Cas9*-expressing eggs after dechorionation only.

	Bleach	Wire-and-tape
No. of eggs	49	23
No. of larvae	24	23
No. of pupae	24	23
No. of adults	24	23
Survivability	**48.9%**	**100%**

Table 4: Survivability of transgenic *D. simulans Cas9*-expressing eggs after dechorionation and injection.

	Bleach and injection	Wire-and-tape and injection
No. of eggs injected	51	16[1]
No. of larvae	11	6
No. of pupae	9	6
No. of adults	9	6
Survivability	**17.6%**	**37.5%**

[1]40 eggs were dechorionated, but due to being put out of proper orientation during that process, only 16 were injected into the posterior ends.

Table 5: Survivability of transgenic *D. melanogaster Cas9*-expressing eggs after being dechorionated with bleach and egg eclosion rate with no dechorionation.

	Bleach dechorionation	No dechorionation
No. of eggs	73	88
No. of larvae	10	44
No. of pupae	10	44
No. of adults	9	44
Survivability	**12%**	**50%**

Table 6: Survivability of transgenic *D. melanogaster Cas9*-expressing eggs when exposed to dechorionation and desiccation/ halocarbon oil or no dechorionation and halocarbon oil.

	Bleach and desiccation only	Bleach and halocarbon oil only	No dechorionation and halocarbon oil only
No. of eggs	64	31	64
No. of larvae	6	3	24
No. of pupae	6	3	24
No. of adults	6	3	20
Survivability	**9.4%**	**9.7%**	**31.2%**

3.2   Transcription of sgRNA

Synthesis of sgRNA for use in targeting genes when injected into eggs can be achieved by either sgRNA plasmids or synthetically transcribing sgRNA. I chose to perform *in vitro* transcription of sgRNA and to inject sgRNA as opposed to using plasmids, as the former has been found to be highly efficient at generating mutations by NHEJ (Bassett & Liu, 2014). This method was used to transcribe sgRNA targets for the four common regions of *fru* in both species (Figure 12), along with additional sgRNA targets for C4 and C5 exons in *D. simulans* and *y* sgRNA (Figure 12) as based on an sgRNA sequence from Bassett & Liu (2014). This process allowed for sgRNA to be prepared in 3 days to use for injections. sgRNAs targeting genes in *D. melanogaster* were then injected into a *D. melanogaster Cas9*-expressing line. A transgenic Cas9 stock was used, as the most efficient CRISPR method involves injecting short synthetic gRNA (sgRNA) into transgenic *Drosophila* embryos expressing *Cas9* (Bassett and Liu*,* 2013; Ren *et al.,* 2013; reviewed in: Housden *et al.,* 2014; Ren *et al.,* 2014). A *D. simulans Cas9*-expressing line was also originally planned to be used for injections, which would first require crossing the *Cas9* into the *D. simulans* FC background. However, reports of weak Cas9 expression in this line (personal communication, Nicholas W VanKuren) meant that it would be more efficient to inject sgRNA and Cas9 directly into *D. simulans* FC instead.
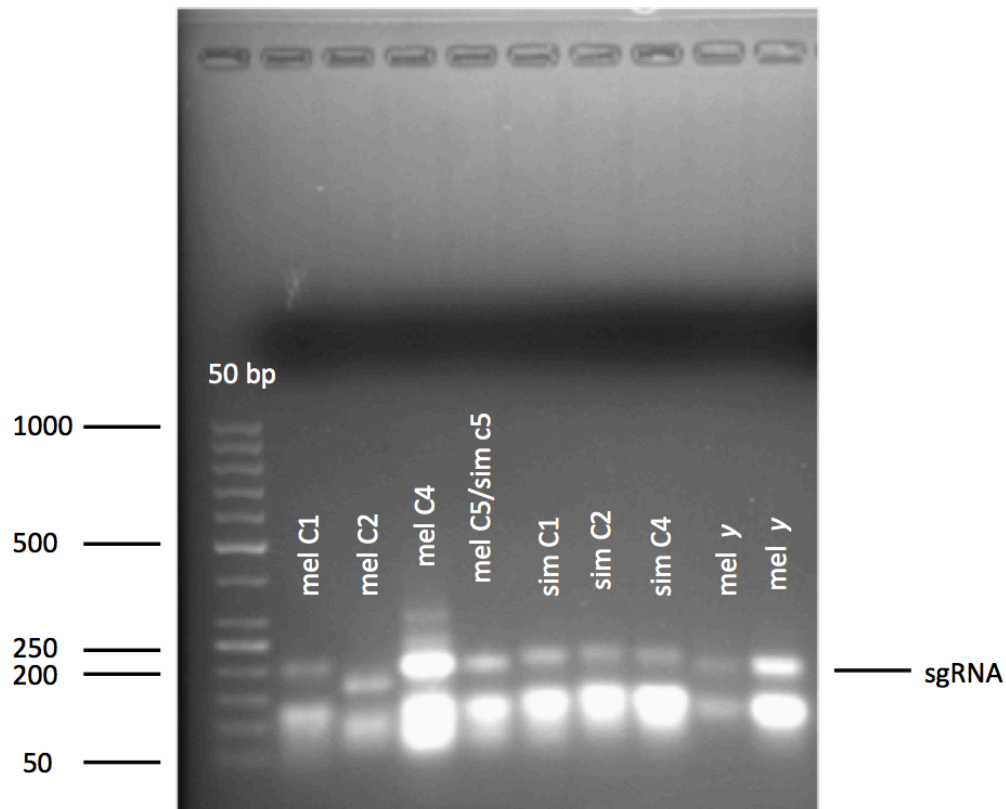
Figure 12: sgRNA transcribed *in vitro* for four *fru* common region exons in *D. melanogaster* (mel) and *D. simulans* (sim) and *y* in *D. melanogaster* (mel).

Each lane received 1 μl of sgRNAs generated for common region exons C1, C2, C4, and C5 of *fru* in *D. melanogaster* and *D. simulans,* and *y* in *D. melanogaster* respectively. The 50bp ladder on the left indicates band sizes for sgRNA. All sgRNAs were ~200 bp in size, with unidentified smearing appearing below this band at ~75bp due to DNA or protein contamination. The extra larger bands for mel C4 could have also been due to contamination, so this particular sgRNA transcript was not used for injections and other purified sgRNA targeting C4 was used.

3.3   Screening injected flies for CRISPR/Cas9 germline mutations

To ensure successful germline mutations of *D. melanogaster* and *D. simulans* with *fru* disruptions, there needed to be an efficient CRISPR method to produce sgRNA and an efficient injection protocol. Injections were first done with *in vitro* transcribed sgRNA targeting *fst* to act as a control and test the efficiency of the injection protocol. These injections were completed at the same time as injections into *D. melanogaster* targeting *fru*. This was followed by injections into *D. melanogaster* targeting *y* as another control to test the injection methods and CRISPR/Cas9-mediated NHEJ efficiency, additional injections into *D. melanogaster* targeting *fru,* and injections into *D. simulans* targeting *fru* common regions C4 and C5.

3.3.1   *D. melanogaster* control injections targeting *fst* and *y*

Control injections were performed to help establish an efficient CRISPR method and injection protocol. Previous *fst* injections by Dr. Anthony Percival-Smith using the *fst* sgRNA vector, the *fst* repair vector, and the Cas9 vector underwent CRISPR-mediated HR with a mutation efficiency of 10% (personal communication, Dr. Anthony Percival-Smith). *In vitro* transcribed *fst* sgRNA and the *fst* repair template were injected into *Cas9*-expressing transgenic *D. melanogaster* to test for recombination efficiency and to act as a control for injections. Germline mutations were screened for a white-eye (w^+) phenotype after mating. Only 15 viable G_0 adults were obtained, none of which produced w^+ progeny (Table 7). This is likely due to low survivability (of 15, one would only expect about 1 individual to have a mutation) and degradation of sgRNA during injections due to use of added food dye in injection mix. Injection mix with food dye was originally used for *fst* injections and *fru* injections into *D. melanogaster* performed by me, as this was how previous *fst* injections were being performed by members of the lab of Dr. Percival-Smith. Food dye was used to help visualize the solution during injections. Food dye was only discovered after injections of *fst* and *fru D. melanogaster* sgRNA to degrade RNA, as it was not tested for RNase contamination during these injections (see below). As there was low survivability and no germline mutations for *fst* injections, this

was not a valuable control for mutation efficiency and to help troubleshoot an efficient injection protocol.

Injections with *y* were added as a second control to test CRISPR-mediated NHEJ germline mutation. Injections of *y in vitro*-transcribed sgRNA injected alongside Cas9 mRNA have been found to have a germline mutation efficiency of 34.5% (Bassett *et al*. 2013, Bassett & Liu, 2014). Using the same design of *y* sgRNA, sgRNA was injected into hybrid transgenic *Cas9*-expressing *D. melanogaster*. These injections were done with and without dechorionation, with both approaches giving low survivability and no germline mutations (Table 7). Therefore, *y* was also not a valuable control to test if *in vitro* transcription of sgRNA is an efficient method. However, it was a valuable control for troubleshooting the use of injection mixes in the injection protocol. During *y* injections, degradation of sgRNA was found to be caused by the food dye used in injection mix. Before and after every round of injections, sgRNA was electrophoresed on a 1.2% TAE gel to check integrity. sgRNA with food dye was electrophoresed after injections and found to be degraded, but sgRNA without food dye was found to be intact (Figure 13). For this reason, all injections following this (injections targeting *y* without dechorionation*, injections performed by Best Gene, and injections targeting C4 and C5 in *D. simulans*) were done without an injection mix and food dye, allowing for sgRNA to maintain its integrity during the injection procedure. As survival with or without dechorionation was relatively the same, the key problem to survival was likely the injection method. However, as I was working with limited injection experience, I did deem the chorion intact method better than the bleach dechorionation method as keeping the chorion intact provided more time to perform injections.

Table 7: Survivability and mutation rates of viable _D. melanogaster_ adults after control injections with or without dechorionation.

	frost **sgRNA and repair template** (Injections with bleach dechorionation & food dye)	_yellow_ (Injections with bleach dechorionation & food dye)	_yellow_ (Injections without dechorionation & without food dye)
No. of eggs injected	205	173	112
No. of larvae	17	7	4
No. of pupae	15	7	4
No. of viable adults	15	4	3
Survivability	**7.8%**	**2%**	**2.7%**
Germline mutations	**0%**	**0%**	**0%**
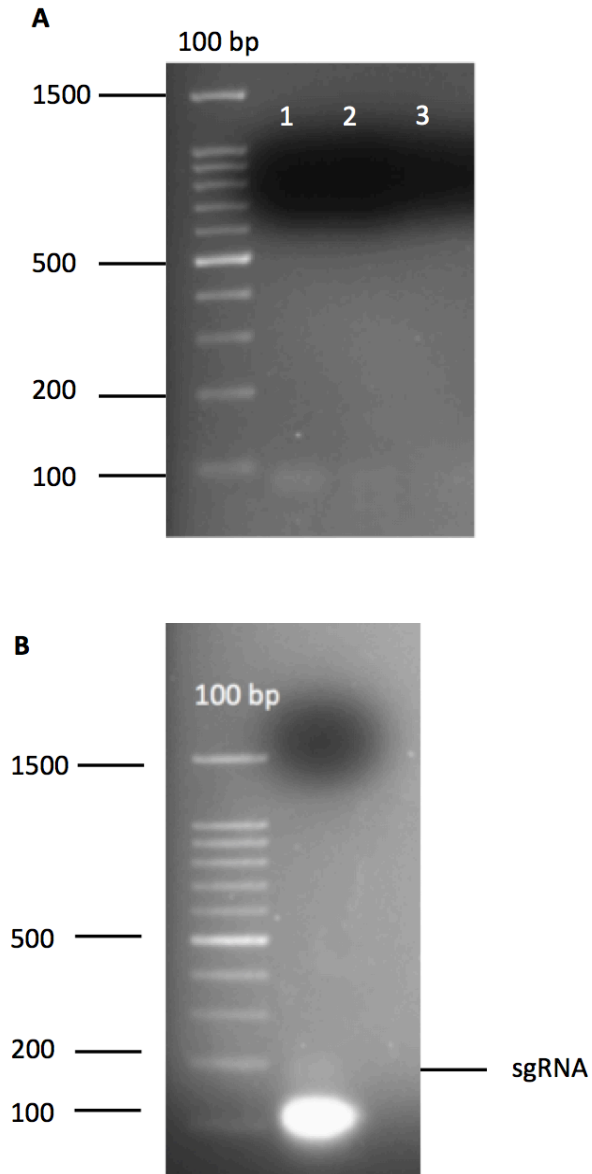
Figure 13: Screening for sgRNA degradation after injections.

1 μl of *yellow* sgRNA ran alongside a 100 bp DNA ladder. **A.** Degradation of sgRNA when mixed with food dye. Lane 1: sgRNA mixed with food dye, kept at -20°C for one hour. Lane 2 and 3: sgRNA mixed with food dye after 4 hours of injections. **B.** sgRNA without food dye is still intact after 4 hours of injections.

3.3.2   *D. melanogaster* injections targeting *fru*

Injections of sgRNA targeting *fru* in *D. melanogaster* transgenic flies expressing *Cas9* were performed such that sgRNA targeting C1 and C2, and sgRNA targeting C4 and C5, were injected together to generate a deletion spanning two exons. The common region exons of *fru* were chosen because removal of them ablates all *fru* functionality, eliminating the need to target *fru* transcripts individually. The first round of injections used C1 and C2 sgRNA or C4 and C5 sgRNA in a mixed solution with food dye into bleach-dechorionated eggs, which gave low adult survivability and no germline mutations (Table 8). Inconsistencies with bleach dechorionation prompted the move to injecting without dechorionation, though injecting sgRNA alongside food dye was still done as it was not yet determined that food dye contributed to the degradation of sgRNA. This approach also resulted in low adult survivability, though more consistent survivability than bleach, as well as no germline mutations (Table 9).

At the same time as I was performing these injections, sgRNA (only sgRNA, no injection mix with food dye) was sent to Best Gene to inject into the same line of transgenic *Cas9*-expressing *D. melanogaster*. This approach resulted in a higher survivability that allowed for screening of a greater amount of viable adults for a deletion in *fru* (Table 10). A large deletion between exons through NHEJ was not achieved, so screening for deletions in only a single exon was done using RE digestion. The idea behind a RE digestion was if a single sgRNA was targeting a sequence as opposed to both sgRNAs, than a smaller deletion in a specific exon could be occurring. If no mutation was present (i.e. the sequence was either not cleaved by Cas9 or was not altered by a repair) then the sequence would be digested by the RE. If the sequence were altered in any way by CRISPR, the sequence would not be cleaved (Figure 10). Deletions were obtained in the C4 exon of *fru* (Figure 14); there was a 4% germline mutation efficiency of C4 using *in vitro* transcribed sgRNA (Table 10). Flies with a *fru* C4 deletion were crossed to *D. melanogaster TM6B* balancer stock with tubby phenotype to maintain the deletion (deletion stocks C4-a, -b, -c). However, stocks with the deletion are homozygous viable. This was unexpected as *fru* deletions are not homozygous viable (Ryner *et al*., 1996).

Table 8: Survivability and mutation rates of viable *D. melanogaster* adults after *fru* injections (dechorionated eggs).

	fru C1/C2 targets (Injections with bleach dechorionation & food dye)	*fru* C4/C5 targets (Injections with bleach dechorionation & food dye)
No. of eggs injected	101	17
No. of larvae	12	0
No. of pupae	8	0
No. of viable adults	5	0
Survivability	**4.9%**	**0%**
Germline mutations	**0%**	**0%**

Table 9: Survivability and mutation rates of viable *D. melanogaster* adults after *fru* injection (no dechorionation).

	fru C1/C2 targets (Injections without bleach dechorionation but with food dye)	*fru* C4/C5 targets (Injections without bleach dechorionation but with food dye)
No. of eggs injected	447	326
No. of larvae	49	20
No. of pupae	47	19
No. of viable adults	37	17
Survivability	**8%**	**5%**
Germline mutations	**0%**	**0%**

Table 10: Survivability and mutation rates of viable *D. melanogaster* adults after *fru* injections (by Best Gene).

	fru C1/C2 targets	*fru* C4/C5 targets
	(Injections without food dye)	(Injections without food dye)
No. of eggs injected	>300	>300
No. of larvae	~90	~70
No. of pupae	89	80
No. of viable adults	89	78
Survivability	**~30%**	**~26%**
Germline mutations	**0%**	**4% in C4 target**

Best Gene provided only approximate numbers for eggs injected and surviving larvae.

A ▼

GCTCAGACCGCTAAGGGCGCTGG

B

Wild-type

CGGCAAGTAATCTGGGGGCAGGGCCACGTGGTGTTCCCGATGCTGCTGCTGCTGCTGGGCGTCCATCAGCT
CCTGCTTGATGCTCAGCGACTGGCTCAGACCGCTAAGGGCGCTGGCCGCGCCCACATTGCCACCACCGCCG
CCCAGATGAACGCTTGTGGTCATCGACAGCGGCGAGTCGTGGCGACCGCCGGCGTCCTCCGATTCCGATGT
GCTGCCGCCCGTTTTGGGCACGGGACTCGACTTGAAGTTGGC

Deletion stock C4-a

CGGCAAGTAATCTGGGGGCAGGGCCACGTGGTGTTCCCGATGCTGCTGCTGCTGCTGGGCGTCCATCAGCT
CCTGCTTGATGCTCAGCGACTGGCTCAGACCGCT---------GGCCGCGCCCACATTGCCACCACCGCCG
CCCAGATGAACGCTTGTGGTCATCGACAGCGGCGAGTCGTGGCGACCGCCGGCGTCCTCCGATTCCGATGT
GCTGCCGCCCGTTTTGGGCACGGGACTCGACTTGAAGTTGGC

Deletion stock C4-b

CGGCAAGTAATCTGGGGGCAGGGCCACGTGGTGTTCCCGATGCTGCTGCTGCTGCTGGGCGTCCATCAGCT
CCTGCTTGATGCTCAGCGACTGGCTCAGACCGC------CGCTGGCCGCGCCCACATTGCCACCACCGCCG
CCCAGATGAACGCTTGTGGTCATCGACAGCGGCGAGTCGTGGCGACCGCCGGCGTCCTCCGATTCCGATGT
GCTGCCGCCCGTTTTGGGCACGGGACTCGACTTGAAGTTGGC

Deletion stock C4-c

CGGCAAGTAATCTGGGGGCAGGGCCACGTGGTGTTCCCGATGCTGCTGCTGCTGCTGGGCGTCCATCAGCT
CCTGCTTGATGCTCAGCGACTGGCTCAGACCGCTAAGGG---TGGCCGCGCCCACATTGCCACCACCGCCG
CCCAGATGAACGCTTGTGGTCATCGACAGCGGCGAGTCGTGGCGACCGCCGGCGTCCTCCGATTCCGATGT
GCTGCCGCCCGTTTTGGGCACGGGACTCGACTTGAAGTTGGC

Figure 14: Germline mutation of *fru* C4 exon in *D. melanogaster*.

A. The sgRNA target sequence (yellow) with the PAM sequence (red) in C4 with the arrowhead marking 3 nucleotides upstream of the PAM sequence where the Cas9 enzyme cleaves the sequence. **B.** The DNA sequence of C4 exon in *D. melanogaster*, with the wild-type sequence at the top, followed by the whole sequence with each deletion obtained (deletion stocks C4-a, -b, -c). Dashes represent where nucleotides in the sequence have been deleted due to DSB and repair by NHEJ.

3.3.3   *D. simulans* injections targeting *fru*

sgRNA was designed for disrupting the same four *D. simulans* common region exons as were targeted in *D melanogaster*. However, due to the above success targeting C4, precedence was taken to attempt first to disrupt C4 and C5 exons in *fru* to generate similar deletions in *D. simulans* as already generated in *D. melanaogaster*. Three additional targets in those two exon sequences were designed and sgRNA for each of these was *in vitro* transcribed; one of these new targets in C4 overlapped the C4 target sequence used in *D. melanogaster* (Table 1, Appendix A). Two sgRNA targeting C4 and two sgRNA targeting C5 were injected together in various combinations, or all sgRNA were injected together. The injection solution (made up of only various sgRNA and Cas9 protein) excluded food dye to prevent sgRNA degradation. This was injected into wild-type *D. simulans* flies; transgenic *D. simulans* flies for *Cas9*-expression were not used due to reports of it being inefficient (personal communication, Nicholas W VanKuren). Injections were done without dechorionation and yielded similar low egg laying and low survivability as seen with *D. melanogaster* injections (Table 10). Injections in *D. simulans* were not successful at generating a large-scale deletion of C4 and C5 sequences, or individual deletions in either exon (Table 11).

Table 11: Survivability and mutation rates of viable *D. simulans* adults after injections targeting *fru*.

	fru C4/C5 targets all sgRNA (Injections without dechorionation & without food dye)	*fru* C4/C5 targets sgRNA A, B[1] (Injections without dechorionation & without food dye)	*fru* C4/C5 targets sgRNA B, C[1] (Injections without dechorionation & without food dye)	*fru* C4/C5 targets sgRNA C, D[1] (Injections without dechorionation & without food dye)
No. of eggs injected	131	162	109	364
No. of larvae	13	12	6	11
No. of pupae	13	11	5	11
No. of viable adults	12	7	3	9
Survivability	**9.2%**	**4.3%**	**2.75%**	**2.5%**
Germline mutations	**0%**	**0%**	**0%**	**0%**

[1]sgRNA target sequences, as found in Table 1.

3.4 Protein analysis of *fru* deletion in *D. melanogaster*

All three different deletion stocks (deletion stocks C4-a, -b, -c) generated through CRISPR/Cas9 in *D. melanogaster* did not result in a frame shift mutation, though deletions at the nucleotide level did generate deletions of, or changes to, amino acids in the Fru protein (Figure 15). Fru protein is known to be a part of the BTB-ZF family, where C1 and C2 of the common region code for the BTB/POZ domain of the protein (Ito *et al.,* 1996). When analyzing the amino acid sequence, Phyre2 did identify this association between C1 and C2 (Figure 16A), but Phyre2 did not identify the C4 amino acid sequence within the BTB/POZ domain, so C4 could not be represented on the protein model. The model also predicted these deletions to have very minor effects on the protein structure as this region of the sequence could tolerate mutations. As the C4 amino acid sequence was not recognized to be part of a conserved domain, this indicates that the deletions I generated in C4 are less likely to affect the functionality of Fru. However, when the C4 amino acid sequence was analyzed alone and not within a Fru protein, the C4 amino acid sequence was shown to form an alpha-helical structure that could have a role in chromosome partitioning, as it was similar to other sequences with this role (Figure 16B).

Wild-type

```
ANFKSSPVPKTGGSTSESEDAGGRHDSPLSMTTSVHLGGGGGNVGAASALSGLSQSLSIKQELMDAQQQQQ
HREHHVALPPDYLP
```

Deletion stock C4-a

```
ANFKSSPVPKTGGSTSESEDAGGRHDSPLSMTTSVHLGGGGGNVGAAS---GLSQSLSIKQELMDAQQQQQ
HREHHVALPPDYLP
```

Deletion stock C4-b

```
ANFKSSPVPKTGGSTSESEDAGGRHDSPLSMTTSVHLGGGGGNVGAASG--GLSQSLSIKQELMDAQQQQQ
HREHHVALPPDYLP
```

Deletion stock C4-c

```
ANFKSSPVPKTGGSTSESEDAGGRHDSPLSMTTSVHLGGGGGNVGAAST-GLSQSLSIKQELMDAQQQQQH
REHHVALPPDYLP
```

Figure 15: Amino acid sequence of C4 protein with the three identified deletions.

The wild-type sequence (top) compared to the amino acid deletions obtained in the three fly stocks of C4 deletions (last three sequences). Dashes indicate where amino acids were deleted in the sequence and red underlined letters indicate an amino acid change. Letters are the short form for their associated amino acid. Deletion lines are as follows: three amino acid (3 a.a.) deletions (deletion stock C4-a), 2 a.a. deletions with 1 a.a. change (deletion stock C4-b), and 1 a.a deletion with 1 a.a. change (deletion stock C4-c).

A　　　　　　　　　　　　　　　　**B**
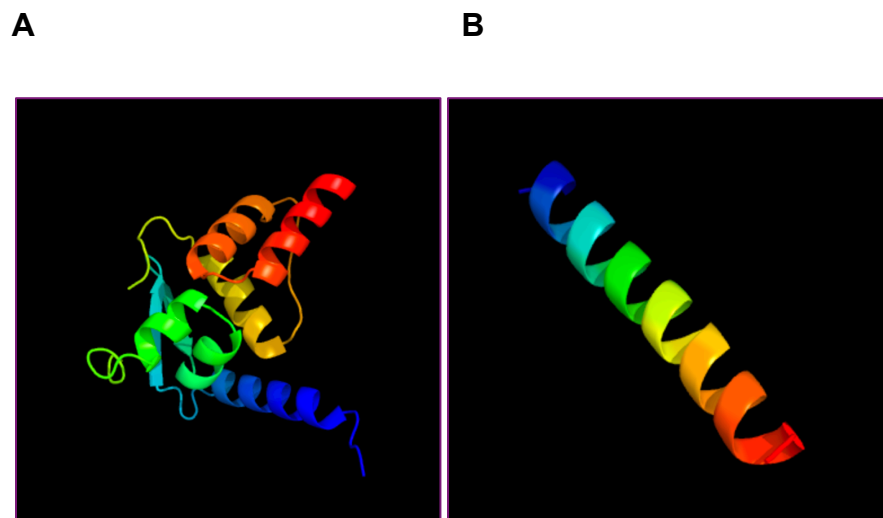
Figure 16: Protein structure of (A) Fru BTB-POZ domain consisting of C1 and C2 sequence, and (B) C4 secondary structure.

Alpha-helix (spiral) and beta (arrow) secondary structures where the rainbow patterning identifies the regions of the protein, with red indicating N terminus and blue indicating C terminus. Protein models designed using Phyre2 (2015). The BTB-POZ domain in Fru (**A**) only consists of the C1 and C2 amino acid sequences so that C3, C4, and C5 sequences were not apart of the modeled structure.

3.5. Conservation of the common region amino acid sequence between *Drosophila* species

I looked at the protein sequences of C4 in *Drosophila* species in order to find out whether the sequence is conserved, suggesting that it is constrained in its evolution in relationship to Fru function, or if it diverges between species, which would suggest divergence of Fru function between species. The amino acid sequence is highly conserved between *D. melanogaster* and *D. simulans,* with identical amino acid sequences. However, there are slight variations in the sequence when compared to species recently diverged from *D. melanogaster* subgroup such as *D. yakuba* (two additional glutamines in the sequence), *D. pseudoobscura* (one alanine deletion, three glycine substitutions, one threonine substitution, and one additional glutamine)*,* and further diverged species such as *D. virilis,* where the sequence is seen to be less conserved (one serine substitution, one threonine deletion, two separate glycine insertions, one arginine substitution, one valine insertion, one valine deletion, two glycine deletions, an additional glutamine, and a glutamic acid substitution) (Figure 17). In these comparisons, species that have longer divergence time have a more highly diverged sequence, as expected. However, the amino acid sequence among these species is identical where the CRISPR/Cas9 deletion was generated (Figure 17). This is interesting because it could mean that this portion of the sequence is constrained in its evolution in relationship to Fru function. As flies with the deletion were homozygous viable (see above) and did not appear to affect Fru structure (Figure 16), the deleted sequence is likely not pivotal for survivability, but the deletion could affect behaviour.

Divergence in sequence amongst species also seems to apply when looking at the amino acid sequences of C3 and C5 (data not shown). In the C5 amino acid sequence, there are two amino acid changes and one deletion between *D. melanogaster* and *D. simulans*, but the sequence becomes even more divergent when compared to *D. yakuba* and *D. pseudoobscura*, with multiple amino acid substitutions and deletions. A homologous amino acid sequence could not be identified for *D. virilis*, or a C5 genomic DNA sequence. C3 was not a focus for targeting deletions in this project, but as part of the

common region, its amino acid sequence was analyzed. C3 amino acid sequences were identical between *D. melanogaster* and *D. simulans*, but had amino acid substitutions and deletions amongst the other species, with the most differences in the sequences observed between *D. pseudoobscura* and *D. virilis*.

The amino acid sequences of C1 and C2, which form the BTB/POZ domain in the zinc-finger protein, are highly conserved among all species examined. The C1 amino acid sequence is identical in all species – except for *D. virilis,* although change in sequence was with two amino acid substitutions. The C2 amino acid sequence is identical in all species examined. The C1 and C2 sequences, composing the BTB/POZ domain, are highly conserved.
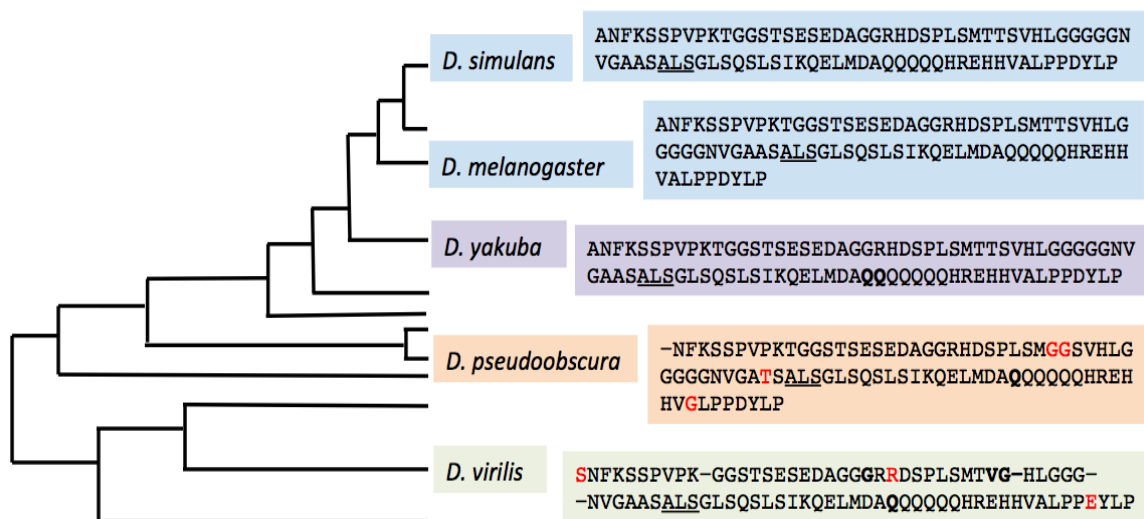
Figure 17: Conservation of C4 amino acid protein sequence between different *Drosophila* species.

A phylogenetic tree showing the divergence of the amino acid sequence of C4 in *fru* between *D. melanogaster, D. simulans, D. yakuba, D. pseudoobscura,* and *D. virilis*. The amino acid sequence is beside each of the represented species, in the same box colour as the species it originates from. Species with the same colour are those that have the same sequence. Amino acid changes in comparison to the *D. melanogaster* sequence are as follows: amino acid insertions are bolded letters, substitutions are red letters, and deletions are represented by dashes. Underlined letters indicate the amino acid changes that were targeted by CRISPR in the three deletion stocks, showing conservation of these amino acids between all species examined.

3.6 Behavioural analysis of *fru* deletion in *D. melanogaster*

Although the amino acid deletions induced by CRISPR do not cause an obvious conformational change in the protein or seem to affect viability, they may still affect the phenotype of the organism. This is because the amino acids that were deleted in C4 were conserved amongst the species, which could mean they have an important role in Fru function, where deletions could disrupt Fru function in behaviour. The third chromosome of these deletion lines are identical to the *Cas9*-expressing *D. melanogaster* stock except that one homolog of the third chromosome has a *fru* mutation, and this third chromosome is maintained over a *TM6B* balancer. The X chromosome and the second chromosome were recombinants, so genotypes of these chromosomes varied. The homozygous males for the *fru* C4 deletion and the Cas9 stock have identical third chromosomes, except for *fru*.

3.6.1 Male behavioural sterility assay

I tested the effect of the deletion on male behaviour since *fru* has been extensively studied in this regard (Hall, 1978; Gailey and Hall, 1989; Burtis and Baker, 1989; Ryner *et al*., 1996; Goodwin *et* al., 2000; Anand *et al*., 2001; Demir and Dickson, 2005; reviewed in: Billeter *et al*., 2006a; Billeter *et al*., 2006b). It is characteristic for males with a fru^M deficiency to be sterile (Gailey and Hall,1989), so I first tested whether males homozygous for the deletion would be sterile due to lack of mating. The deletion C4-a did not cause for males to become sterile as they copulated frequently with females and produced offspring. Chaining behaviour, a characteristic for fru^M deficient males, where males form a line and court each other, was also not observed in the deletion line amongst males. However, increased mating in the deletion line compared to the wild-type control males was observed (Table 12). 94% of mutant males produced offspring within the mating assay, while 68% of controls produced offspring (z-test, P= 0.00174, N = 50). The significantly low mating in control males compared to *fru* deletion line was unexpected, as controls are expected to mate at 100% efficiency as seen in the sterility assay performed by Gailey and Hall (1989). The control males used were transgenic *D. melanogaster* flies expressing *Cas9*, and have been observed to be slow at mating during this project (data not shown).

Table 12: Behavioural phenotype of homozygous *fru* C4-a mutant males compared to wild-type males.

Genotype	Mating fraction
fru C4-a deletion	47/50
Wild-type	34/50

3.6.2   Female behavioural mating assay

I also tested the effect of the deletion on female hybrids' rejection of *D. melanogaster* males to see if the C4-a deletion line could be used to generate hybrids for the reciprocal hemizygosity test. Courtship and copulation in *D. melanogaster* controls and hybrids was scored based on observation in the first hour of the behavioural assay. Copulation in the hybrids was scored on presence or absence of sperm in the female reproductive tract 24 hours after the initial behaviour assay, as hybrids often take longer to mate then the controls. C4-a deletion mutants/*D. simulans* female hybrids are heterospecific, except at the *fru* locus, where the *D. melanogaster* allele contains the CRISPR-induced deletion. These females showed high levels of mating, equivalent to both the control (balancer) hybrids and *D. melanogaster* pure species females when paired with *D. melanogaster* males (G-test, P = 0.848, N = 30; Figure 18). This test did not match results from previous behavioural mating assays, where *fru* disruptions saw significantly reduced mating in the hybrid *fru* disruption females (Moehring lab, unpublished data). As a result, the C4 deletion lines could not be used for the reciprocal hemizgosity test, as the deletion did not affect female preference.
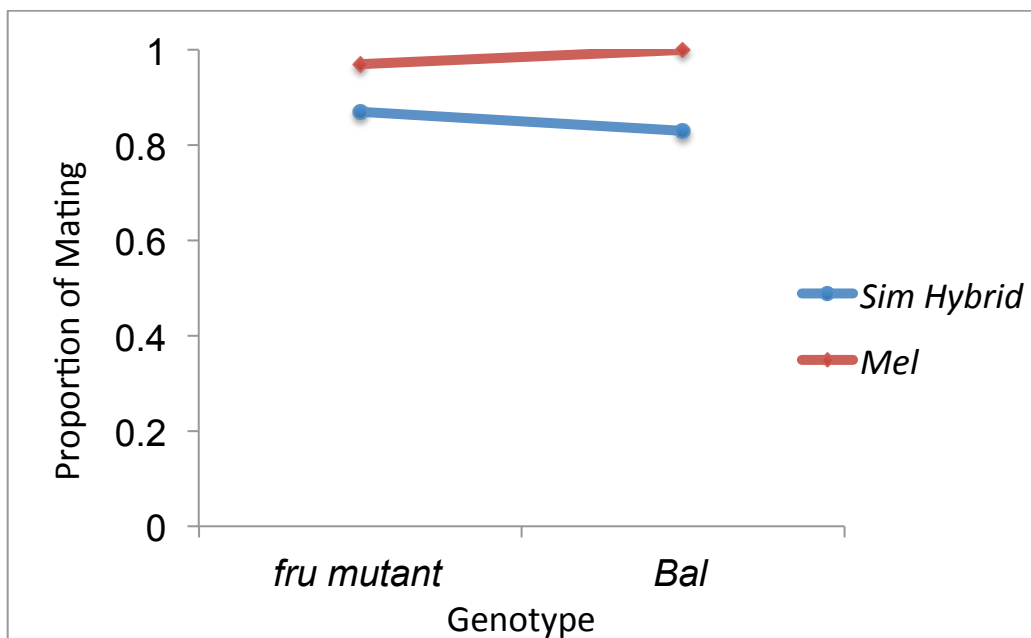
Figure 18: Mating proportions of hybrid females with *fru* mutation to wild-type *D. melanogaster* males compared to controls.

The blue line represents *D. melanogaster*/*D. simulans* hybrid females (*Sim* Hybrid) with *fru* mutation or *Bal* genotype, the red line represents *D. melanogaster* females (*Mel*) with *fru* mutation or *Bal* genotype. All females were scored for courtship and copulation with *D. melanogaster* males to see if hybrid females with *fru* mutation would show reduced mating compared to the other three genotypes, due to unmasking of the *D. simulans fru* allele. No significant change in mating (P=0.848, N=30) for the hybrid mutant was observed.

4   Discussion

4.1 Efficiency and optimization

A component of this project was to find a protocol that could be easily used for generating sgRNA to target *fru* and allow for disruption of the targeted exon in *fru*. In addition, a working protocol had to be established for microinjection, which would allow for efficient injection of CRISPR components into fly embryos for germline mutations.

4.1.1   CRISPR

CRISPR/Cas9 has been used for genome manipulation since 2013, making it a new technique that is still being developed (reviewed in: Wu *et al*., 2014). Its first use was in mammalian cells to test the efficiency of injecting Cas9 and sgRNA vectors to target a sequence of interest and generate a DSB to be repaired by either NHEJ (Cong *et al*., 2013) or HR (Mali *et al*., 2013), both of which were effective.

In regards to generating mutations using NHEJ, as this project aimed to do, four different methods have been explored in *D. melanogaster* and proven to be efficient at targeting *y* or *w*. The first method required the injection of two plasmids into embryos: one plasmid for Cas9 and one for sgRNA. The Cas9 plasmid expressed the *Cas9* gene under *Hsp70* promoter, and the sgRNA plasmid was the expression of sgRNA driven by the *U6* promoter: a recognized highly efficient promoter for expressing sgRNA (Gratz *et al*., 2013; reviewed in: Bassett and Liu, 2013). The second method also required the injection of sgRNA and Cas9 into embryos, but this method used *in vitro* transcribed sgRNA under the T7 promoter and *in vitro* transcribed *Cas9* expressed either through the T7 or Sp6 promoter. This method had higher mutagenesis rates and thus a greater efficiency (reviewed in: Bassett and Liu, 2013; Yu *et al*., 2013; Bassett, *et al*., 2013; Bassett and Liu, 2014). The third method was a transgenic approach where flies that transgenically expressed *Cas9* under the *nanos* promoter were crossed to transgenic flies expressing sgRNA under *U6* promoter (Kondo and Ueda, 2013; reviewed in: Bassett and Liu, 2013). This method is the most efficient as it can produce more than 90% mutagenic flies, but is the most time consuming because it is more difficult molecularly to generate two

transgenic lines, and requires extra generations of crossing flies. The fourth and final method takes strategies from the first and third methods: plasmids encoding sgRNA are injected in *Cas9*-expressing transgenic flies where *Cas9* is expressed under the *vasa* or *nanos* promoters (Ren *et al*., 2013; Sebo *et al.,* 2014; reviewed in: Bassett and Liu, 2013). This method is thought to be the best compromise of all of the methods outlined in terms of mutation efficiency, effort, and time required (reviewed in: Housden *et al*., 2014).

An alternate method uses parameters of techniques two and four: *in vitro* transcribed sgRNA is injected into transgenic *Cas9* flies. This approach had not yet been tested for efficiency (reviewed in: Bassett and Liu, 2013), but was employed in this project and was able to produce mutations in the C4 exon of *fru*, demonstrating that this is another viable technique to use in *Drosophila*. However, this exact method was only applicable for use in *D. melanogaster* as a *D. simulans* stock with efficient expression of Cas9 has not been acquired. Use of the CRISPR/Cas9 system in *D. simulans* still involved *in vitro* transcribed sgRNA, but sgRNA was co-injected with Cas9 protein into wild-type flies. This strategy can be seen as an effective solution as direct injection of Cas9 protein into mammalian cells has been found to reduce the amount of off-target effects compared to plasmid injections (Ramakrishna *et al*., 2014; Kim *et al*., 2014; reviewed in: Wu *et al*., 2014).

As NHEJ germline mutation was achieved in *D. melanogaster* at a rate of 4%, this new technique can be used to generate mutations, but is not overly efficient, as other CRISPR-mediated NHEJ germline mutation efficiencies in *D. melanogaster* have ranged from 4%-88% (Bassett *et al*., 2013). This could be due to the sgRNAs used and the target gene, as some sgRNAs are more effective than others in targeting the same gene (Bassett and Liu, 2014). Injection of multiple sgRNAs targeting different sequences in the same gene can be used as a way to overcome this, as efficiency seems to rely on the sgRNA sequence and DNA sequence being targeted (reviewed in: Wu *et al*., 2013). Multiple designs of sgRNA sequences targeting other sequences in each common region exon of *D. simulans* were designed to overcome the problem of only getting cleavage of C4 exon when DSBs in two exons were desirable.

Transcription of sgRNA was relatively fast, with *in vitro* transcription taking no more than a maximum of three days and sgRNA could be directly injected into flies once made. Though *in vitro* transcription is quick, a limited amount of sgRNA is made, and so more sgRNA has to be repeatedly transcribed if many injections are performed.

4.1.2   Microinjections

Setting up effective microinjection within the lab proved to be a large hurdle and still needs to be improved to have success in creating transformed flies with CRISPR/Cas9. One set-back was collecting enough eggs every 30 minutes for injections. A single female can lay 50-70 eggs in one day, and so a fly cage with many females can be expected to produce over 100 eggs every 30 minutes (Tyler, 2000). The aim is to inject 50-100 eggs every 30 minutes (personal communication, Dr. Anthony Percival-Smith). All flies were kept in store-bought large and small fly cages (Genesee Scientific) that were placed in incubators. When using transgenic *D. melanogaster* flies, only 1-30 eggs would be collected every 30 minutes, making it difficult to inject a large amount of eggs. A reason for this could have been the males of transgenic *D. melanogaster Cas9*-expressing line that were used, as they seemed slow at courtship and copulation. When injecting to target *y*, using wild-type *D. melanogaster* males alongside *Cas9*-expressing *D. melanogaster* females did increase the number of eggs being laid in the chambers, such that 15-30 eggs were likely to be collected every 30 minutes, but egg laying was still low. Using wild-type *D. simulans* in fly cages resulted in a slight increase in eggs laid each half hour, being 15-60 eggs, which allowed for a greater number of egg injections to be performed, and a closer number to the anticipated 50 eggs, which is the largest amount I was capable of injecting every 30 minutes. This low egg yield could have affected the lack of CRISPR/Cas9 transformed flies, as injecting so few eggs and having a low survivability after injection lowered the odds of how many flies could have a targeted deletion.

To overcome the problem of low egg yield, I employed such strategies as crowding the fly cage with more flies, using younger flies (those that were under a week old), changing the agar juice plates constantly the day before injection, changing the cage every two weeks to ensure that the chamber was not dirty, and making grooves in the agar juice

plate to help induce egg laying. None of these approaches dramatically increased egg yield, though they did help to bring up the number of eggs laid into the range of 15-60 eggs as listed above. One factor that was found to affect egg yield was weather, specifically change in barometric pressure. It was observed that on warmer, sunny days, more eggs were laid than on stormy days when the pressure was lower. Another factor that could be affecting egg yield is the noise of the incubator where the cages are stored, as noise does reduce mating and egg laying (personal communication, Dr. Anne Simon). Therefore it is recommended that fly cages be kept outside of incubators in a temperature and humidity controlled room, and to maintain a flexible schedule when it comes to injecting, such that injections can take place when the barometric pressure is ideal for egg laying. The circadian rhythm of flies should also be considered, so that injections can take place during peak egg laying times. Additionally, collecting eggs from more than one fly cage could help. During this project, eggs were collected from only one or two cages, but additional cages could be made out of polypropylene beakers or plastic cups and covered with fine stainless steel or nylon mesh to cut costs.

Another aspect of microinjections that needed to be troubleshot was whether injecting dechorionated eggs or eggs with their chorion intact would be a better option. One option that was tested once and quickly discarded due to lack of access to a vacuum where the fly cages were kept was dechorionation with a filtration apparatus (Sullivan *et al*., 2000; Cartwright, 2009). Two other methods that I explored (see Results chapter), was the use of bleach or wire-and-tape for removing the chorion. Once again mechanics of dechorionation is what gave bleach an advantage and made it the best method of dechorionation for this project. However, dechorionation still gave low survivability and so the avenue of injecting with the chorion intact was explored. Keeping the chorion intact did not significantly change the survival of eggs compared to dechorionation. Therefore, the issue experienced with low survivability was likely due to the microinjection process, as very few larvae would emerge from the eggs that were injected.

To explore this problem, further survivability tests were done to look at desiccation of the eggs and the use of halocarbon oil (Table 6). For eggs that were dechorionated,

desiccation and covering eggs in halocarbon oil only decreased survivability by <3%, while eggs intact, covered in oil, decreased in survivability by ~19%, though this could have been due to random variation. However, this does showcase the problem of leaving halocarbon oil on the eggs after injection; halocarbon oil must be drained off of the eggs after injection (Gompel and Schröder, 2005; Bassett and Liu, 2014), which was always ensured at time of injections in this project. Additionally, it is recommended to aerate the halocarbon oil before injections (Gompel and Schröder, 2005), though this was not done in this project but is recommended for future tests.

These tests did not fully explain the lethality issues being experienced, and so other problems must be due to injections themselves. These issues could be: overloading the egg with injection mix, putting the needle too far into the egg, and having excessive leakage of cytoplasm (Sullivan *et al.*, 2000; Gompel and Schröder, 2005; Cartwright, 2009). To try and overcome these issues, the pressure of the needle was adjusted to only release a small bubble of injection mix, a very thin needle tip was used, as larger tips caused greater leakage (Sullivan *et al.*, 2000; Cartwright, 2009), and when dechorionating the egg, desiccating the egg for the correct amount of time helped to prevent leakage (Sullivan *et al.*, 2000)

In regards to whether injecting with or without a chorion, it seems that the choice is down to personal preference and equipment available. Keeping the chorion intact does make it more difficult to visualize sgRNA entering the egg, but halocarbon oil or use of ethanol can help to make the chorion more translucent. My experience testing survivorship with each method does show survivorship is dependent on a number of different factors in the microinjection protocol. If a needle puller is available to pull needles appropriately, then one can inject with the chorion intact (Sullivan *et al.*, 2000 p. 354). Additionally, the lesser amount of disturbance to the eggs caused by skipping the dechorionation process is regarded by some to be preferable (personal communication with Dr. Graeme Maiden of Genetivision).

4.2   Targeting deletions of *fru*

In both *D. melanogaster* and *D. simulans*, targeted deletions of *fru* were aimed to cause DSBs in the common region exons such that a portion of the sequence of two exons could be deleted through NHEJ repair. It was also critical for deletions to occur in the germline so that a stock of *D. melanogaster* and *D. simulans* with a *fru* disruption could be maintained and used to generate hybrids with a single species-specific allele of *fru.*

4.2.1   CRISPR-mediated NHEJ in *D. melanogaster*

Injections of sgRNA targeting four exons of the common region in *D. melanogaster* were done by me and through Best Gene. Although a large sequence deletion between two exons was not obtained in either case, a sequence deletion in the C4 exon was achieved. Three different germline deletions in C4 (Figure 14) all resulted in the loss of amino acids (Figure 15), although there is likely no change to protein structure from these deletions (Figure 16).

As the C4 amino acid sequence is present in all Fru proteins, it was expected that this deletion would affect the expression of *fru*. A characteristic of *fru* mutants is homozygous lethality (Ryner *et al*., 1996). However, individuals homozygous for the C4 deletion are viable, indicating that the deletion did not affect survivability. Though this is the case, there was still the possibility for the mutation to be affecting behaviour. This was explored in the *fru* deletion stock with three amino acid deletions: deletion stock C4-a (as this stock was stable at the time) by performing sterility and behavioural mating assays to see the potential effect on male and female mating behaviour (Table 12, Figure 18).

The deletion in C4 did not inhibit males courtship or copulation abilities as seen in other *fru* mutants (Gailey and Hall, 1989), though only null Fru^{M} mutations have been found to abolish male courting behaviour completely. All other Fru^{M} mutant males still retain the ability to court females, whether the females were conspecific or heterospecific (Fan *et al*., 2013). Thus, I can conclude that the C4 deletion does not completely disrupt Fru^{M} expression. However, there was a significant increase in copulation due to the mutation. The higher rate of copulation in the mutant compared to the control was not seen before

between *fru* mutants and wild-type flies (Gailey and Hall, 1989). These mutants and the control share the same third chromosome, except for *fru*, but the X chromosome and second chromosome were not identical between the mutants and control, which could account for the differences in mating. It would therefore be ideal to study this change in behaviour more closely, by performing the mating assay again with wild-type *D. melanogaster* flies that do not carry Cas9, and with the mutation crossed into the same wild-type background, to determine if the mutation still produces a higher copulation success. In addition, male courtship behaviour in the deletion line could be quantified to see how much time the male spends in each of the courtship steps and if the timing is different from that seen in the control.

The deletion also did not significantly reduce mating of the hybrid females with the deletion when paired with *D. melanogaster* males, which would have been expected if the deletion allowed for the recessive expression of the *D. simulans* allele that affected female preference (Laturney and Moehring, 2012a; Moehring lab, unpublished data). These findings could indicate that perhaps the deletion was too small to affect the molecular function of non-sex-specific Fru proteins, as it was not capable of disrupting Fru^M expression. Alternatively, perhaps the C4 exon is not pivotal for *fru*'s role in female behaviour, and the sequence changes can be tolerated so that the function of *fru* in behaviour remains unaffected. It could however hold another role in females that is not behaviour-specific, such as a developmental role.

Only the C1 and C2 exons contribute to the BTB/POZ domain, (Zollman *et al.*, 1994; Ito *et al.*, 1996), a major component of the BTB-zinc finger (BTB-ZF) motif of Fru, acting as a protein-protein interaction motif that mediates transcriptional regulation (Yamamoto, *et al.,* 2004; Stogios, *et al.*, 2005). When looking into the conservation of the common region across different *Drosophila* species, the amino acid sequence of the BTB/POZ domain was highly conserved, while the other sequences of the common region, namely C4 and C5, had more sequence divergence, though the sequence where the deletion was obtained was conserved. This indicates that there is some divergence in the common region, especially in exons outside of the BTB/POZ domain, but overall the region is mostly conserved amongst *Drosophila*. However, as the C4 sequence was conserved

between closely related species (*D. melanogaster* and *D. simulans*), but had more sequence changes in *Drosophila* species outside of that clade, perhaps it does contribute to speciation through molecular divergence of Fru between species. Previous work has used the *fru* DNA sequence, particularly the BTB-domain of *fru,* as a way to classify different *Drosophila* species (Gailey *et al*., 2000), as BTB domain families have undergone lineage specific expansions, and so can be used to classify different species (Stogios, *et al*., 2005). This indicates that divergence in the common region may play a role in species diversification, but the role of C4 is still unknown. If C4 does not play a behavioural role, it could have another role in females that is yet determined. Perhaps it is connected more to female development than behaviour, as non sex-specific isoforms that have C4 are involved with early development (Ryner *et al*., 1996; Anand *et al*., 2001; Song *et al*., 2002; Dornan *et al*., 2005; Neville *et al*., 2014). With the divergence in C4 sequence between distantly related species, we could expect for species with large variations in the common region sequences to have different Fru protein isoforms, change the efficiency and timing of protein binding during neuronal development, or have Fru interact with different proteins. For example, this later case is seen in *D. yakuba* and *D. suzukii,* where Tra does not suppress the female-specific *fru* transcript expression as it does in other *Drosophila* species (Yamamoto *et al*., 2004). As the C4 sequence is identical in *D. melanogaster* and *D. simulans*, it could have a role that is evolutionary conserved, such as an effect on female development that went unnoticed in this project. Alternatively, perhaps C4 does not have a large role to play in protein function, which is why there is divergence in sequence between species and it is not as highly conserved as C1 and C2 exons. This could mean that the deletions do not greatly affect Fru function, and as such are not integral to gene expression.

Since C4 does code for an amino acid sequence that is seen in all Fru isoforms, it would be beneficial to find out what role C4 has in the Fru protein. As C3, C4, and C5 are in between the BTB domain and the zinc-finger (ZF) domains, they could form the middle linker region; a region observed in some of the most common BTB families that follows the BTB domain (Stogios *et al*., 2005). In BTB-ZF proteins, the linker region has been seen to interact with accessory proteins to aid in chromatin remodeling and transcription repression, and its sequence is often not as highly conserved as the BTB or ZF regions

(Stogios *et al*., 2005). Therefore, C4 inclusion in the protein could be beneficial for correct folding and spacing of the protein to help form protein binding sites (Ito *et al*., 1996), though this could not be modeled (Figure 16). Or, as these sequences are present in all protein variants of *fru,* even where ZF motifs are not present (Yamamoto *et al*., 2004), perhaps they act as a tether for the BTB domain to bind to proteins, such as in aiding the formation of the Fru-Bon complex (Ito *et al*., 2012). Perhaps all of the common region exons outside of the BTB domain are required for the association of HDAC1 or HP1a with the Fru-Bon complex, which drives or suppresses masculinization of neurons (Ito *et al*., 2012). This in turn could help in the formation of neurons dictating mating behaviour. As such, there may be a molecular evolution of *fru* whereby the modification of the common region could dictate the protein complex formation for Fru with other proteins, which could dictate the activity of the Fru zinc-finger isoforms and control the transcription factor activity and efficiency of Fru on downstream genes in neurons responsible for mating behaviour.

4.2.2   CRISPR-mediated NHEJ in *D. simulans*

With the success of a deletion in the C4 exon of *fru* in *D. melanogaster*, the aim for ensuring a sequence deletion in *D. simulans* took precedence for targeting C4 and C5. However, due to lack of targeting two exons in the common region to cause a large sequence deletion in *D. melanogaster*, the strategy of injecting multiple sgRNA targeting C4 and C5, as opposed to one target for each exon, was employed. When injecting with multiple sgRNAs, different combinations of sgRNAs were used, such as injecting all eight sgRNAs, or injecting four sgRNAs (two sgRNAs for each exon).

No deletion was obtained, though this is likely due to low survivability after injection. If the *D. melanogaster* deletion in C4 was only obtained at a 4% germline mutagenesis rate, then low mutation efficiency could also be possible for *D. simulans*. But with only less than 9% of injected embryos surviving to become viable adults, it would not have been likely to get a survivor with a mutation, as observed.

4.3 Limitations

As a summary of the discussion above, there were limitations in my study in the approaches used, the fly lines used, and the portion of the gene being targeted. Firstly, there was a large learning curve for microinjections as I had to establish a protocol for my lab. I was unable to produce a highly efficient protocol, as there were problems with low survivability after injection, which affected my ability to disrupt *fru*, but a protocol is in place that can be improved upon. I was also using a relatively new methodology to generate CRISPR products by injecting *in vitro* transcribed sgRNA into transgenic Cas9 *D. melanogaster* flies. The effectiveness of this protocol in *Drosophila* has not yet been reported (reviewed in: Basset and Liu, 2013). sgRNA transcription produced a blob or smear below the sgRNA band (Figure 12). The cause of this is unknown, but could have been due to leftover reagents from the transcription protocol that were not properly separated from sgRNA during purification. Though deletions in C4 were still obtained from sgRNA that had a smear, this could have affected the efficiency of sgRNA targeting. Another reason why efficiency could have been low was insufficient Cas9 expression in *D. melanogaster* transgenic flies. Insufficient expression of Cas9 was seen in transgenic *Cas9*-expressing *D. simulans* flies, which caused modification of the injection protocol; I injected Cas9 along with sgRNA, which could have also had an effect on achieving germline mutations.

Though deletions in *fru* were obtained in *D. melanogaster*, the deletion was only obtained in a single exon. It has been stated earlier that target efficiency relies on both the sgRNA sequence and the DNA target, with some sgRNA being more effective than others (reviewed in: Wu *et al.*, 2013). However, it is also possible that the DNA sequences of the common region are not effective targets for sgRNA. CRISPR deletions of *fru* are currently only available for the P exons and A-C exons (fly lines available from Dr. Stephen Goodwin)*,* but there are no reports of CRISPR-mediated deletions in the common region.

Additionally, all the protocols tested for efficient CRISPR germline mutation have currently only been done in *D. melanogaster* and efficiency of each technique is not reported in other *Drosophila* species. As *D. simulans* is closely related to *D.*

melanogaster it was assumed that any CRISPR technique used would be similar in efficiency for *D. simulans*. However, the efficiency of CRISPR-mediated NHEJ germline mutations varies across different species, from 5.5% efficiency in mosquitoes (Dong *et al*., 2015) to 40%-60% efficiency in beetles (Gilles *et al*., 2015) to 70%-100% efficiency in mice (Li, *et al*., 2013; Qin, *et al*., 2015), showcasing how CRISPR as a tool can be better utilized in some species than in others.

4.4 Future directions

Due to difficulties in getting mutations in *fru* through CRISPR-mediated NHEJ in both species, this project could not reach the end goal of performing a reciprocal hemizygosity test, to test female species-specific mating behaviour amongst hybrids with *fru* disruptions. As a result, the focus turned to generating *fru* disruption lines that could be used in the future to test behaviour.

Though a deletion in *fru* in *D. melanogaster* was generated through targeting the common region of *fru*, this deletion did not have an effect on female behaviour, as described above, and so a *fru* deletion in *D. melanogaster* will need to be attempted again. This can be achieved by designing additional sgRNA targets for the common region, just as was done for *D. simulans*. Focus on generating a larger deletion by causing DSBs in two exons will likely cause disruption of *fru*. Ideally, targeting C1 and C2 exons may be able to disrupt *fru* as these exons give rise to the BTB domain of Fru (Ito *et al*., 1996). Another way to increase chances of disrupting *fru* would be to improve upon the microinjection protocol used here, by troubleshooting with injections targeting *y* to help increase egg survivability. Then, following the method of *in vitro* transcription of sgRNA and microinjection described here, successful mutations can be generated. Additionally, with no successful *D. simulans fru* disruption, disruption lines will also have to be made, ideally targeting the same exons as those that are targeted in *D. melanogaster* to ensure the deletion can be identical in both species to overcome any off-target effects.

As genotyping for deletions caused by NHEJ was more time consuming due to PCR and RE digests, it might be beneficial to focus on generating mutations through alternative approaches. Focusing on CRISPR-mediated HR, a visual marker could be inserted into

fru where the two DSBs are made in the common region. This will allow for easier scoring to look for *fru* disruptions. Taking this approach may require changing the CRISPR protocol, as injecting *in vitro* transcribed sgRNA alongside Cas9 or injecting into transgenic flies has not been tested for efficiency with HR, just NHEJ (reviewed in: Bassett and Liu, 2013). Therefore it is recommended that for CRISPR-mediated HR, the approach that can be taken is to inject the donor template and plasmids containing the sgRNA sequence expressed under the *U6b* promoter (Port *et al*., 2014) into transgenic flies expressing *Cas9,* as this approach is effective for time and effort (reviewed in: Housden *et al*., 2014). Alternatively, traditional methods of designing CRISPR can be replaced: recently, scientists have been using t-RNA flanked sgRNAs. These have been recognized to be highly efficient for targeting when working with multiple sgRNA targets (Port and Bullock, 2016).

Once disruptions of *fru* can be obtained in both *D. melanogaster* and *D. simulans*, they can be crossed to the opposite species to generate hybrids that only express a species-specific allele of *fru*. These female hybrids can then be tested for female mating preference through a reciprocal hemizygosity test by pairing with wild-type *D. melanogaster* males. The protocol to perform a reciprocal hemizygosity test can be found in Appendix C.

4.5   Conclusion

Establishing disruptions of *fru* in both *D. melanogaster* and *D. simulans* through CRISPR-mediated NHEJ was troublesome and had low efficiency. As a germline deletion was obtained in *D. melanogaster*, targeting the common region through *in vitro* transcribed sgRNA injected into transgenic flies expressing *Cas9* is possible. Therefore, the initial steps of forming a working CRISPR-mediated NHEJ protocol and a microinjection protocol has been established upon which the future steps of this project can be completed. However, if someone with injection experience does not continue this project, injections could be outsourced to a company such as Best Gene, which produced the C4 deletion stocks in *D. melanogaster*.

I generated stable stocks of *fru* C4 mutations in *D. melanogaster* that can be used for future experiments to understand what molecular effect, if any, the deletion has on the function of *fru*, giving new insight into the role of the common region. This information can further help in unraveling the story of why *D. melanogaster* and *D. simulans* are reproductively isolated from each other.

References

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Gaile, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., Chen, L. X., Brandon, R. C., Rogers, Y. H., Blazej, R. G., Champe, M., Pfeiffer, B. D., Wan, K. H., Doyle, C., Bater, E. G., Helt, G., Nelson, C. R., Gabor, G. L., Abril, J. F., Agbayani, A., An, H. J., Andrews-Pfannkoch, C., Baldwin, D., Ballew, R. M., Basu, A., Baxendale, J., Bayraktarogy, L., Beasley, E. M., Beeson, K. Y., Benos, P. V., Berman, B. P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M. R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K. C., Busam, D. A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J. M., Cawley, S., Dahlke, C., Davenport, L. B., Davies, P., de Pablos, B., Delcher, A., Deng, Z., Mays, A. D., Dew, I., Dietz, S. M., Dodson, K., Doup, L. E., Downes, M., Dugan-Rocha, S., Dunkov, B. C., Dunn, P., Durbin, K. J., Evangelista, C. C., Ferraz, C., Ferriera, S., Fleischmann, W., Fosler, C., Gabrielian, A. E., Garg, N. S., Gelbart, W. M., Glasser, K., Glodek, A., Gong, F., Gorrell, J. H., Gu, Z., Guan, P., Harris, M., Harris, N. L., Harvey, D., Heiman, T. J., Hernandez, J. R., Houck, J., Hostin, D., Houston, K. A., Howland, T. J., Wei, M. H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G. H., Ke, Z., Kennison, J. A., Ketchum, K. A., Kimmel, B. E., Kodira, C. D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A. A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T. C., McLeod, M. P., McPherson, D., Merkulov, G., Milshina, B., Mobarry, C., Morris, J., Moshrefi, A., Mount, S. M., Moy, M., Murphy, B., Murphy, L., Muzny, D. M., Nelson, D. L., Nelson, D. R., Nelson, K. A., Nixon, K., Nusskern, D. R., Pacleb, J. M., Palazzolo, M., Pittman, G. S., Pan, S., Pollard, J., Puri, V., Reese, M. G., Reinert, K., Remington, K., Saunders, R. D., Scheeler, F., Shen, H., Shue, B. C., Siden-Kiamos, I., Simpson, M., Skupski, M P., Smith, T., Spier, E., Spradling, A. C., Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R., Venter, E., Wang, A. H., Wang, X., Wang, Z. Y., Wassarman, D. A., Weinstock, G. M., Weissenbach, J., Williams, S. M., Woodage, T., Worley, K. C., Wu, D., Yang, S.,

Yao, Q. A., Ye, J., Yeh, R. F., Zaveri, J. S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X. H., Zhong, F. N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H. O., Gibbs, R. A., Myers, E. W., Rubin, G. M., & Venter, J. C. (2000). The genome sequence of *Drosophila melanogaster*. *Science, 287,* 2185-2195.

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

Anand, A, Villella, A., Ryner, L. C., Carlo, T., Goodwin, S. F., Song, H. J., Gailey, D. A., Morales, A., Hall, J. C., Baker, B. S., Taylor, B. J. (2001). Molecular genetic dissection of sex-specific and vital functions of the *Drosophila melanogaster* sex determination gene *fruitless*. *Genetics, 158*, 1569-1595.

Arias, A. M. (2008). *Drosophila* Methods and Protocols. *Drosophila melanogaster* and the development of biology in the 20th century. *Methods in Molecular Biology*, *420*, 1-25.

Baker, B. S., Taylor, B. J., & Hall, J. C. (2001). Are complex behaviors specified by dedicated regulatory genes? Reasoning from *Drosophila*. *Cell, 105,* 13-24.

Bassett, A. R. & Liu, J. L. (2013). CRISPR/Cas9 and genome editing in *Drosophila*. *Journal of Genetics and Genomics, 41*, 7-19.

Bassett, A. R. & Liu, J. L. (2014). CRISPR/Cas9 mediated genome engineering in *Drosophila*. *Methods, 69,* 128-136.

Bassett, A. R., Tibbit, C., Ponting, C. P., & Liu, J. L. (2013). Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Reports, 4*, 220-228.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., & Nekrasov, V. (2013). Plant genome editing made easy: targeted mutagenesis in model and crop plants using CRISPR/Cas system. *Plant Methods, 9*, 1-10.

Bell, L. R., Horabin, J. I., Schedl, P., Cline, T. W. (1991). Positive autoregulation of *Sex-lethal* by alternative splicing maintains the female determined state in *Drosophila. Cell, 65,* 229-239.

Belote, J. M. & Baker, B. S. (1982). Sex determination in *Drosophila melanogaster*: analysis of *transformer-2*, a sex-transforming locus. *Proceedings of the National Academy of Sciences, 79,* 1568-1572.

Belote, J. M., McKeown, M., Boggs, R. T., Ohkawa, R., Sosnowski, B. A. (1989). Molecular genetics of *transformer*, a genetic switch controlling sexual differentiation in *Drosophila*. *Genesis, 10,* 143-154.

Beumer, K. J. & Carroll, D. (2014). Targeted genome engineering techniques in *Drosophila, Methods, 68,* 29-37.

Bhaya, D., Davidson, M., & Barrangou, R. (2011). CRISPR-Cas system in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annual Review in Genetics, 45,* 273-297.

Billeter, J. C., Rideout, E. J., Dornan, A. J., Goodwin, S. F. (2006a). Control of male sexual behavior in *Drosophila* by the sex determination pathway. *Current Biology, 16,* R766-R776.

Billeter, J. C., Villella, A., Allendorfer, J. B., Dornan, A. J., Richardson, M., Gailey, D. A., & Goodwin, S. F. (2006b). Isoform-specific control of male neuronal differentiation and behavior in *Drosophila* by the *fruitless* gene. *Current Biology, 16,* 1063-1076.

Brambilla, M., Vitulano, S., Spina, F., Baccetti, N., Gargallo, G., Fabbri, E., GuiDali, F., & Randi, E. (2008). A molecular phylogeny of the *Sylvia cantillans* complex: Cryptic species within the Mediterranean basin. *Molecular Phylogenetics and Evolution, 48,* 461-472.

Burtis, K. C. & Baker, B. S. (1989). *Drosophila doublesex* gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell, 56*, 997-1010.

Butlin, R. K. & Ritchie, M. G. (1989). Genetic coupling in mate recognition systems: what is the evidence? *Biological Journal of the Linnean Society, 37,* 237-246.

Cande, J., Stern, D. L., Morita, T., Prud'homme, B., & Gompel, N. (2014). Looking under the lamp post: neither *fruitless* nor *doublesex* has evolved to generate divergent male courtship in *Drosophila. Cell Reports, 8,* 363-370.

Capy, P., Pla, E., & David, J. R. (1993). Phenotypic and genetic variability of morphometrical traits in natural populations of *Drosophila melanogaster* and *D. simulans*. I. Geographic variations. *Genetics Selection Evolution, 25*, 517-536.

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

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

Cartwright, E. J. (2009). Transgenesis Techniques. *Methods in Molecular Biology, 561*, 3-19.

Chung, H., Loehlin, D. W., Dufour, H. D., Vaccarro, K., Millar, J. G., & Carroll, S. B. (2014). A single gene affects both ecological divergence and mate choice in *Drosophila*. *Science, 343,* 1148-1151.

Clynen, E., Ciudad, L., Belles, X., & Piulachs, M. D. (2011). Conservation of Fruitless' role as master regulator of male courtship behaviour from cockroaches to flies. *Development Genes and Evolution*, *221*, 43-48.

Cobb, M., Connolly, K., Burnet, B. (1985). Courtship behaviour in the *Melanogaster* species sub-group of *Drosophila*. *Behaviour, 95,* 203-230.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science, 339*, 819-823.

Coyne, J. A. & Orr, H. A. (1989). Patterns of speciation in *Drosophila*. *Evolution, 43,* 362-381.

Coyne, J. A. & Orr, H. A. (2004). Speciation. Massachusetts: Sinaur Associates Press.

Demir, E. & Dickson, B. J. (2005). *fruitless* splicing specifies male courtship behavior in *Drosophila*. *Cell, 121*, 785-794.

Deveau, H., Barrangou, R., Garneau, J. E., Labonte, J., Fremaux, C., Boyaval, P. Romero, D. A., Horvath, P., & Moineau, S. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *Journal of Bacteriology, 190*, 1390-1400.

DiCarlo, J. E., Norville, J. E., Mali, P., Rios, X., Aach, J., & Chruch, G. M. (2013). Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas9 systems. *Nucleic Acids Research, 41*, 4336-4343.

Ding, Q., Regan, S. N., Xia, Y., Oostrom, L. A., Cowan, C. A., & Musunuru, K. (2013). Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell, 12*, 393-394.

Dobzhansky, T. (1937). Genetics and the Origin of Species. New York: Columbia University Press.

Dong, S., Lin, J., Held, N. L., Clem, R. J., Passarelli, A. L., & Franz, A. W. E. (2015). Heritable CRISPR/Cas9-mediated genome editing in the yellow fever mosquito, *Aedes aegypti*. *PLoS One, 10*, e0122353.

Dornan, A. J., Gailey, D. A., & Goodwin, S. F. (2005). GAL4 enhancer trap targeting of the *Drosophila* sex determining gene *fruitless*. *Genesis, 42*, 236-246.

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

Finck, J., Kuntze, J., & Ronacher, B. (2016). Chemical cues from females trigger male courtship behaviour in grasshoppers. *Journal of Comparative Physiology A, 202,* 337-345.

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

Finley, K. D., Taylor, B. J., Milstein, M., & McKeown, M. (1997). *dissastisfaction*, a gene involved in sex-specific behaviour and neural development of *Drosophila melanogaster. Proceedings of the National Academy of Sciences, 94*, 913-918.

Gailey, D. A., Billeter, J. C., Liu, J. H., Bauzon, F., Allendorfer, J. B., & Goodwin, S. F. (2006). Functional conservation of the *fruitless* male sex-determination gene across 25 myr of insect evolution. *Molecular Biology and Evolution, 23*, 633-643.

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

Gailey, D. A., Taylor, B. J., & Hall, J. C. (1991). Elements of the *fruitless* locus regulates development of the muscle of Lawrence, a male-specific structure in the abdomen of *Drosophila melanogaster* adults. *Development, 113*, 879-890.

Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A. H., & Moineau, S. (2010). The

CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature, 468,* 67-71.

Gilles, A. F., Schinko, J. B., & Averof, M. (2015). Efficient CRISPR-mediated gene targeting and transgene replacement in the beetle *Tribolium castaneum. Development, 142*, 2832-2839.

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

Gramates, L. S., Marygold, S. J., dos Santos, G., Urbano, J. M., Antonazzo, G., Matthews, B. B., Rey, A. J., Tabone, C. J., Crosby, M. A., Emmert, D. B., Falls, K., Goodman, J. L., Hu, Y., Ponting, L., Schroeder, A. J., Strelets, V. B., Thurmond, J., Zhou, P., & the FlyBase Consortium (2017). FlyBase at 25: looking to the future. *Nucleic Acids Research, 45*, D663-D671.

Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison, M. M., Wildonger, J., & O'Conner-Giles, K. M. (2013). Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics, 194*, 1029-1035.

Gratz, S. J., Ukken, F. R., Rubinstein, C. D., Thiede, G., Donohue, L. K., Cummings, A. K., & O'Connor-Giles, K M. (2014). Highly specific and efficient CRISPR/Cas9-catayzed-homology-directed repair in *Drosophila. Genetics, 196*, 961-971.

Greenspan, R. & Ferveur, J. (2000). Courtship in *Drosophila. Annual Review in Genetics, 34*, 205-232.

Gompel, N. & Schröder, E. A. (2005). *Drosophila* germline transformation protocol. http://gompel.org/wp-content/uploads/2015/12/Drosophila-transformation-with-chorion.pdf

Goodwin, S. F., Taylor, B. J., Villella, A., Foss, M., Ryner, L. C., Baker, B. C., & Hall, J. C. (2000). Aberrant splicing and altered expression patterns in *fruitless* mutants of *Drosophila melanogaster. Genetics, 154,* 725-745.

Haber, J. E. (1995). *In vivo* biochemistry: physical monitoring of recombination induced by site-specific endonucleases. *Bioessays, 17,* 609-620.

Haft, D. H., Selengut, J., Mongodin, E. F., & Nelson, K. E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Computational Biology, 1*, e60.

Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L., Terns, R. M., & Terns, M. P. (2009). RNA-guided RNA cleavage by CRISPR RNA-Cas protein complex. *Cell, 139,* 945-956.

Hall, J. C. (1978). Courtship among males due to a male-sterile mutation in *Drosophila melanogaster. Behavior Genetics,* 8, 125-141.

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

Heinrichs, V., Ryner, L. C. & Baker, B. S. (1998). Regulation of sex-specific selection of *fruitless* 5' splice sites by *transformer* and *transformer-2. Molecular and Cellular Biology, 18*, 450-458.

Housden, B. E., Lin, S., & Perrimon, N. (2014). Cas9-based genome editing in *Drosophila. Methods in Enzymology, 546,* 415-439.

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

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J. R. J., & Joung, J. K. (2013). Efficient genome engineering in zebrafish suing CRISPR-Cas system. *Nature Biotechnology, 31*, 227-229.

Ito, H., Fujitani, K., Usui, K., Shimizu-Nishikawa, K., Tanaka, S., & Yamamoto, D. (1996). Sexual orientation in *Drosophila* is altered by the satori mutation in the sex-determination gene *fruitless* that encodes a zinc finger protein with a BTB domain. *Proceedings of the National Academy of Sciences, 93*, 9687-9692.

Ito, H. Sato, K., Koganezawa, M., Ote, M., Matsumoto, K., Hama, C., & Yamamoto, D. (2012). Fruitless recruits two antagonistic chromatin factors to establish single-neuron sexual dimorphism. *Cell, 149,* 1327-1338.

Jakociunas, T., Bonde, I., Herrgard, M., Harrison, S. J., Kristensen, M, Pederson, L. E., Jensen, M. K., & Keasling, J. D. (2015). Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae. Metabolic Engineering, 28*, 213-222.

Jansen, R., van Embden, J. D. A., Gaastra, W., & Schoulis, M. N. (2002). Identification of genes associated with DNA repeats in prokaryotes. *Molecular Microbiology, 43*, 1565-1575.

Jao, L. E., Wente, S. R., & Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proceedings of the Natural Academy of Sciences, 110*, 13904-13909.

Jennings, B. H. (2011). *Drosophila* − a versatile model in biology and medicine. *Materials Today, 14*, 190-195.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science, 337,* 816-821.3

Juni, N. & Yamamoto, D. (2009). Genetic analysis of *chaste*, a new mutation of *Drosophila melanogaster* characterized by extremely low female sexual receptivity. *Journal of Neurogenetics, 23*, 329-340.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction, and analysis. *Nature Protocols, 10*, 845-858.

Kim, S., Kim, D., Cho, S. W., Kim, J., & Kim, J. S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonuclease. *Genome Research, 24*, 1012-1019.

Kimura, K., Hachiya, T., Koganezawa, M., Tazawa, T., & Yamamoto, D. (2008). *fruitless* and *doublesex* coordinate to generate male-specific neurons that can initiate courtship. *Neuron, 59,* 759-769.

Kofler, R., Hill, T., Nolte, V., Betancourt, A. J., Schlotter, C. (2015). The recent invasion of natural *Drosophila simulans* populations by the *P*-elements. *Proceedings of the National Acdemy of Sciences, 112*, 6659-6663.

Kohler, R. E. (1994). Lord of the fly: *Drosophila* genetics and the experimental life. Chicago: The University of Chicago Press.

Kondo, S. & Ueda, R. (2013). Highly improved gene targeting by germ-line specific Cas9 expression in *Drosophila*. *Genetics, 195*, 715-721.

Konopka, R. J. & Benzer, S. (1971). Clock mutants of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, 68*, 2112-2116.

Lachaise, D., Cariou, M. L., David, J. R., Lemeunier, F., Tsacas, L., & Ashburner, M. (1988). Historical biogeography of *Drosophila melanogaster* species subgroup. *Evolutionary Biology, 22*, 159-225.

Laturney, M. & Moehring, A. J. (2012a). Fine-scale genetic analysis of species-specific female preference in *Drosophila simulans*. *Journal of Evolutionary Biology, 25,* 1718-1731.

Laturney, M. & Moehring, A. J. (2012b). The genetic basis of female mate preference and species isolation in *Drosophila*. *International Journal of Evolutionary Biology, 2012*, 1-13.

Lee, G., Foss, M., Goodwin, S. F., Carlo, T., Taylor, B. J., Hall, J. C. (2000). Spatial, temporal, and sexually dimorphic expression patterns of the *fruitless* gene in the *Drosophila* nervous system. *Journal of Neurobiology, 43;* 404-426.

Li, D., Qiu, Z., Shao, Y., Chen, Y., Guan, Y., Liu, M., Li., Y., Gao, N., Wang, L., Lu, X., Zhao, Y., & Liu, M. (2013). Heritable gene targeting in the mouse and rat using CRISPR-Cas system. *Nature Biotechnology, 31*, 681-683.

Lieber, M. R., Ma. Y., Pannicke, U., & Schwarz, K. (2003). Mechanism and regulation of human non-homologous DNA end-joining. *Nature Reviews Molecular Cell Biology, 4,* 712-720.

Liou, W. L. & Price, T. D. (1994). Speciation by reinforcement of premating isolation. *Evolution, 48*, 1451-1459.

Majerus, M., O'Donald, P., & Weir, J. (1982). Evidence for preferential mating in *Adalia bipunctata. Entomology Monthly Magazine, 133*, 55-61.

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I., Yakunin, A. F., van der Oost, J., & Koonin, E. V. (2011). Evolution and classification of the CRISPR-Cas system. *Nature Reviews Microbiology, 9;* 467-477.

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science, 339*, 823-826.

Marraffini, L. A. & Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science, 322*, 1843-1845.

Mayr, E. (1942). Systemics and the Origin of Species. New York: Columbia University Press.

McVey, M., Radut, D., & Sekelsky, J. J. (2004). End-joining repair of double-strand breaks in *Drosophila melanogaster* is largely DNA ligase IV independent. *Genetics, 168*, 2067-2076.

Moehring, A. J. & Mackay, T. F. C. (2004). The quantitative genetic basis of male mating behaviour in *Drosophila melanogaster*. *Genetics, 167,* 1249-1263.

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

Moulin, B., Aubin, T., Jallon, J. M. (2004). Why there is one-way crossability between *D. melanogaster* and *D. simulans*. *Genetica*, *120*, 285-292.

Moulin, B., Rybak, F., Aubin, T., Jallon, J. M. (2001). Compared ontogenesis of courtship song components of males from the sibling species, *D. melanogaster* and *D. simulans*. *Behavior Genetics, 31*, 299-308.

Nagoshi, R. N., McKeown, M., Burtis, K. C., Belote, J. M., & Baker, B. S. (1988). The control of alternative splicing at genes regulating sexual differentiation in *D. melanogaster*. *Cell*, *53*, 229-236.

Nanda, P. & Singh, B. N. (2012). Behavioural reproductive isolation and speciation in *Drosophila*. *Journal of Biosciences, 37,* 359-374.

Neville, M. C., Nojima, T., Ashley, E., Parker, D. J., Walker, J., Southall, T., Van de Sande, B., Marques, A. C., Fischer, B., Brand, A. H., Russell, S., Ritchie, M. G., Aerts, S., & Goodwin, S. F. (2014). Male-specific *fruitless* isoforms target neurodeveleopmental genes to specify a sexually dimorphic nervous system. *Current Biology, 24,* 229-241.

O'Kane, C. J. & Asztalos, Z. (1999). Sexual behaviour: Courting *dissatisfaction. Current Biology, 9*, R289-R292.

Panda, S., Hogenesch, J. B., & Kay, S. A. (2002). Circadian rhythms from flies to humans. *Nature, 417*, 329-335.

Parks, A. L., Cook, K. R., Belvin, M., Dompe, N. A., Fawcett, R., Huppert, K., Tan, L. R., Winter, C. G., Bogart, K. P., Deal J. E., Deal-Herr, M. E., Grant, D., Marcinko, M., Miyazaki, W. Y., Robertson, S., Shaw, K. J., Tabios, M., Vysotskaia, V., Zhao, L., Andrade, R. S., Edgar, K. A., Howie, E., Killpack, K., Milash, B., Norton, A., Thao, D., Whittaker, K., Winner, M. A., Friedman, L., Margolis, J., Singer, M. A., Kopczynski, C., Curtis, D., Kaufman, T. C., Plowman, G. D., Duyk, G., & Francis-Lang, H. L. (2004). Systematic generation of high-resolution deletion coverage of the *Drosophila melanogaster* genome. *Nature Genetics, 36*, 288-292.

Pasyukova, E. G., Viera, C., & Mackay, T. F. C. (2000). Deficiency mapping of quantitative trait loci affecting longevity in *Drosophila melanogaster. Genetics, 156,* 1129-1146.

Port, F. & Bullock, S. L. (2016). Augmenting CRISPR applications in *Drosophila* with t-RNA flanked sgRNAs. *Nature Methods, 13*, 852-854.

Port, F., Chen, H. M., Lee, T., & Bullock, S. L. (2014). Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila. Proceedings of the National Academy of Sciences, 111,* 2967-2976.

Qin, W., Dion, S. L., Kutny, P. M., Zhang, Y., Cheng, A. W., Jillete, N. L., Malhotra, A., Geurts, A. M., Chen, Y. G., & Wang, H. (2015). Efficient CRISPR/Cas9-mediated genome editing in mice by zygote electroporation of nuclease. *Genetics*, *200*, 423-430.

Ramakrishna, S., Kwaku Dad, A. B., Beloor, J., Gopalappa, R., Lee, S. K., & Kim, H. (2014). Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Research, 24,* 1020-1027.

Ren, X., Sun, J., Housden, B. E., Hu, Y., Rosesel, C., Lin, S., Liu, L. P., Yang, Z., Mao, D., Sun, L., Wu, Q., Ji, J., Mohr, S. E., Xu, J., Perrimon, N., & Ni, J. Q. (2013). Optimized gene editing technology for *Drosophila melanogaster* using germ-line specific Cas9. *Proceedings of the National Academy of Sciences, 110,* 19012-19017.

Ren, X., Yang, Z., Li, J. B., & Ni, J. Q. (2014). Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in *Drosophila. Cell Reports, 9*, 1151-1182.

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

Ryder, E. & Russel, S. (2003). Transposable elements as tools for genomics and genetics in *Drosophila. Briefings in Functional Genomics and Proteomics, 2,* 57-71.

Ryner, L. C., Goodwin, S. F., Castrillon, D. H., Anand, A., Villella, A., Baker, B. S., Hall, J. C., Taylor, B. J., & Wasserman, S. A. (1996). Control of male sexual behavior and sexual orientation in *Drosophila* by the *fruitless* gene. *Cell, 87,* 1079-1089.

Schaefer, K. A., Wu, W. H., Colgan, D. F., Tsang, S. H., Bassuk, A. G., & Mahajan, V. B. (2017). Unexpected mutations after CRISPR-Cas9 editing *in vivo*. *Nature Methods*, *14*, 547-548.

Schutt, C. & Nothinger, R. (2000). Structure, function and evolution of sex-determining systems in *Dipteran* insects. *Development, 127*, 667-677.

Sebo, Z. L., Lee, H. B., Peng, Y., & Guo, Y. (2014). A simplified and efficient germline-specific CRISPR/Cas9 system for *Drosophila* genomic engineering. *Fly, 8,* 52-57.

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J. J., Qiu, J. L., & Gao, C. (2013). Targeted genome modification of plant crops using CRISPR-Cas9 system. *Nature Biotechnology, 31*, 686-688.

Shaw, K. L. (1996). Polygenic inheritance of a behavioral phenotype: interspecific genetics of song in the Hawaiian cricket genus *Laupala*. *Evolution, 50*, 256-266.

Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., & Siksnys, V. (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO Journal, 30*, 1335-1342.

Sokolowski, M. B. (2001). *Drosophila*: genetics meets behaviour. *Nature Reviews, 2*, 879-890.

Song, H. J., Billeter, J. C., Reynaud, E., Carlo, T., Spana, E. P., Perrimon, N., Goodwin, S. F., Baker, B. S., & Taylor, B. J. (2002). The *fruitless* gene is required for the proper formation of axonal tracts in the embryonic central nervous system of *Drosophila*. *Genetics, 162,* 1703-1724.

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

Spradling, A. C. & Rubin, G. M. (1982). Transposition of cloned *P* elements into *Drosophila* germ line chromosomes. *Science, 218*, 341-347.

Stern, D. L. (2014). Identification of loci that cause phenotypic variation in diverse species with the reciprocal hemizygosity test. *Trends in Genetics, 30*, 547-554.

Stockinger, P., Kvitsiani, D., Rotkopf, S., Tirian, L., & Dickson, B. J. (2005). Neural circuitry that governs *Drosophila* male courtship behavior. *Cell, 121*, 795-807.

Stogios, P. J., Downs, G. S., Jauhal, J. J. S., Nandra, S. K., & Prive, G. G. (2005). Sequence and structural analysis of BTB domain proteins. *Genome Biology, 6*, R82.

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

Sullivan, W., Ashburner, M., & Hawley, R. S. (2000). *Drosophila* protocols. New York: Cold Spring Harbor.

Suzuki, K., Juni, N., Yamamoto, D. (1997). Enhanced mate refusal in female *Drosophila* induced by a mutation in the *spinster* locus. *Applied Entomology and Zoology, 32,* 235-243.

Tauber, E., Roe, H., Costa, R., Hennessy, J. M., & Kyricou, C. P. (2003). Temporal mating isolation driven by a behavioural gene in *Drosophila, Current Biology, 13*, 140-145.

Ting, C.T., Takahashi, A., Wu, C. I. (2001). Incipient speciation by sexual isolation in *Drosophila*: Concurrent evolution at multiple loci. *Proceedings of the National Academy of Sciences, 98*, 6709-6713.

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

Tyler, M. S. (2000). Developmental biology, a guide for experimental study. Sunderland: Sinauer Associates, Inc.

Usui-Aoki, K., Ito, H., Ui-Tei, K., Takahashi, K., Lukacsovich, T., Awano, W., Nakata, H., Piao, Z. F., Nilsson, E. E., Tomida, J., & Yamamoto, D. (2000). Formation of the male-specific muscle in female *Drosophila* by ectopic *fruitless* expression. *Natural Cell Biology, 2*, 500-506.

Vedenina, V. Y., Panyutin, A. K., & von Helversen, O. (2007). The unusual inheritance pattern of the courtship songs in closely related grasshopper species of the *Chorthippus albomarginatus*-group. *European Society for Evolutionary Biology, 20*, 260-270.

Venken, K. J. T. & Bellen, H. J. (2005). Emerging technologies for gene manipulation in *Drosophila melanogaster*. *Nature Review Genetics, 6*, 167-178.

Villella, A., Gailey, D. A., Berwald, B., Ohshima, S., Barnes, P. T., & Hall, J. C. (1997). Extended reproductive roles of the *fruitless* gene in *Drosophila melanogaster* revealed by behavioral analysis of new *fru* mutants. *Genetics, 147*, 1107-1130.

Vincze, T., Posfai, J., & Roberts, R. J. (2003). NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Research, 31*, 3688-3691.

von Philipsborn, A. C., Jorchel, S., Tirian, L., Demir, E., Morita, T., Stern, D. L., & Dickson, B. J. (2014). Cellular and behavioural functions of *fruitless* isoforms in *Drosophila* courtship. *Current Biology, 24,* 242-251.

Von Schilcher, F. (1976). The function of pulse song and sine song in the courtship of *Drosophila melanogaster, Animal Behaviour, 24*, 622-625.

Wang, Y., Li, Z., Xu, J., Zeng, B., & Ling, L. (2013). The CRISPR/Cas9 system mediates efficient genome engineering in *Bombyx mori. Cell Research, 23*, 1414-1416.

Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., Zhang, F., & Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas9-mediated genome engineering. *Cell, 153,* 910-918.

Watanabe, T. K. & Kawnishi, M. (1979). Mating preference and the direction of evolution in *Drosophila, 205,* 906-907.

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

Wiedenheft, B., Sternberg, S. H., & Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature, 482,* 331-338.

Wiernasz, D. C. & Kingsolver, J. G. (1992). Wing melanin pattern mediates species recognition in *Pieris occidentalis. Animal Behavior, 43,* 89-94.

Wu, X., Kriz, A. J., & Sharp, P. A. (2014). Target specificity of the CRISPR/Cas9 system. *Quantitative Biology, 2,* 59-70.

Wu, Y., Liang, D., Wang, Y., Bai, M., Tang, W., Bao, S., Yan, Z., Li, D., & Li, J. (2013). Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell Stem Cell, 13*, 659-662.

Yamamoto, D., Usui-Aoki, K., & Shima, S. (2004). Male-specific expression of the Fruitless protein is not common to all *Drosophila* species. *Genetica, 120*, 267-272.

Yu, Z., Ren, M., Wang, Z., Zhang, B., Rong, Y. S., Jiao, R., & Gao, G. (2013). Highly efficient genome modifications mediated by CRISPR/Cas9 in *Drosophila*. *Genetics, 195*, 289-291.

Zollman, S., Godt, D., Prive, G. G., Couderc, J. L., & Laski, F. A. (1994). The BTB domain, found primarily in zinc finger proteins, defines an evolutionary conserved family that includes several developmentally regulated genes in *Drosophila. Proceedings of National Academy of Sciences, 91,* 10717-10721.

Zordon, M. A. & Sandrelli, F. (2015). Circadian clock dysfunction and psychiatric disease: could fruit flies have a say? *Frontiers in Neurology, 6*, e2015.

Zouros, E. (1981). The chromosomal basis of sexual isolation in two sibling species of *Drosophila: D. arizonensis* and *D. mojavensis. Genetics, 97*, 703-718.

Appendices

Appendix A: Common region sequences

Annotated sequences of the *fru* common region exons C1, C2, C4, and C5 in lines *D. melanogaster* BJS and *D. simulans* FC containing the sgRNA targets. The *D. melanogaster* and *D. simulans* sequences are reverse compliments to each other.

sgRNA target site used

Alternate sgRNA target site that could be used in future study

D. melanogaster

C1

CTTGACTGTTTCGCCCTCGCAGGCGAGCGTGAACGTCGCATAGCGCCTCCCGC
TGCAGCAGTGAGGTTAGCACGCCGGTCAAATTTGTGGGATGATTGTTCCAGC
GCAAGCAGAATTGCTGGTCCATCGCTCCTTGGTCAGTGTTGTAC

The reverse complement of this sgRNA site was used for targeting.

C2

CTGCAGGCTCTCGGCCGTCTTGAGAAACATGGGCAGCGAACTCTGGCCCACG
TTGACCTCGCCCTTGTACATGAAGTCGAGCAGAGATCGCATCTCTGAGTATCT
GACATCTTTCAAGTAAGATGATGGGATGTGGATGCTGGTTCTGTAGGAAAAT
CGTCTCGAAGTACGGACTGCAGGCTGACAGGATGGTCTGGTGAGC

C4

CGGCAAGTAATCTGGGGGCAGGGCCACGTGGTGTTCCCGATGCTGCTGCTGC
TGCTGGGCGTCCATCAGCTCCTGCTTGATGCTCAGCGACTGGCTCAGACCGCT
AAGGGCGCTGGCCGCGCCCACATTGCCACCACCGCCGCCCAGCTGAACGCTT

GTGGTCATCGACAGCGGCGAGTCGTGGCGACCGCCGGCGTCCTCCGATTCCGATGTGCTGCCGCCCGTTTTGGGCACGGGACTCGACTTGAAGTTGGC

C5

CAAAATAATTATATTCTTTTATTAATGCTTCTCAGTTTATTGTTCAATTTTCGC
CAATGTCAAATGCTGTTTGCTTTTGTTCATAGTAGCTGCCAAGTTCTTAGGTT
ACTAAATATATTTTACTTTTCGGACATTTAATTGATAACGAAAACTTAATGCT
AATGCCTATGGCTAATCCTGAGCGGTGTGTTGGGTCCTCTTTCAAGGACCTGT
CTCTATGCGCTGTTCTTGCCCACATTTATGGTGATAATTTCGGGGGAACGTCG
CTTGGAGCTGCCTCCTGCACTTCCACTGGGCATGGGATGTTGCTGCTGCTCCT
GGAGCTCACGCTCGTAGCACATGCGATGCGATTTGCATTTGTTGTGGTGGTTG
ACGATCTCCAGGCGAAAGTTGGAGTTGCCCTCAACGAAGGAGGTGCGTATGT
CGGCCAGCTCGCCGCTCCATTTGCGACAATGGCGGCAGAACATGGTGTTGCC
CGCTCGTCGTATTGCAGCCACTCGAACTGATACAGCCAGTTGGCCCTAAATC
GACCACTTTTCCGGGAGCTTCCTCCCGCAGAGCCCGATGGTATCGCTGGGCTT
TGGGATTTTCCCAGCGAATGGTGTGGCTGCTGTTGATGCTGTTGTTGCTGCTC
CTGGTGTTGTTGCTGCTGTGCCAGCAATCTGTGCATGGCAGAATCCAAGCCCT
CGTTCTGATAATGCAGCTGAAAGCTCTTCAACTGAGCCTCCACTGCCTGGTGG
AGGAGCATTTGATGCATGGACTTGTTGCCTGTGCTGCCTAAAAACGCTTCGTG
ACTTTTTCAC

D. simulans

C1

GTACAACACTGACCAAGGAGCGATGGACCAGCAATTCTGCTTGCGCTGGAAC
AATCATCCCACAAATTTGACCGGCGTGCTCACCTCACTGCTGCAGCGGGAGG
CGCTATGCGACGTCACGCTCGCCTGCGAGGGCGAAACAGTCAAG

C2

GCTCACCAGACCATCCTGTCAGCCTGCAGTCCGTACTTCGAGACGATTTTCCT
ACAGAACCAGCATCCACATCCCATCATCTACTTGAAAGATGTCAGATACTCA
GAGATGCGATCTCTGCTCGACTTCATGTACAAGGGCGAGGTCAACGTGGGTC
AGAGTTCGCTGCCCATGTTTCTCAAGACGGCCGAGAGCCTGCAG

C4

GCCAACTCCAAGTCGAGTCCCGTGCCCAAAACCGGCGGCAGCACATCGGAGT
CGGAGGACGCCGGCGGTCGCCACGATTCGCCGCTCTCGATGACCACCAGCGT
TCATCTGGGCGGCGGTGGTGGCAATGTGGGCGCGGCCAGCGCCCTGAGCGGT
CTGAGCCAGTCGCTGAGCATCAAGCAGGAGCTGATGGACGCCCAGCAGCAG
CAGCAGCATCGGGAACACCACGTGGCCCTGCCCCCAGATTACTTGCCG

C5

GTGAAAAAGTCAGAAGCGTTTTTAGGCAGCACTGGCAACAAGTCCATGCACC
AAATGCTCCTCCACCAGGCAGTGGAGGCTCAGTTGAAGAGCTTTCAGCTGCA
TTACCAGAACGAGGGCTTGGATTCCGCCATGCACAGATTGCTGGCACAGCAG
CAACACCAGGAGCAGCAACAGCAGCACCAGCAGCAGCCACACCTTTCGCTG
GGAAAATCCCAAAGCCCAGTGATACCATCGGGCTCTGCGGAGGAAGCTCCCG
GAAAAGTGGTCGATTTAGGGCAACTGGCTGTATCAGTTCGAGTGGCTGCAAT
ACGACGAGCGGGCCAACACCATGTTCTGCCGCCACTGTCGCAAATGGAGCGG
TGAGCTGGCCGACATACGCACCTCCTTCGTGGAGGGCAACTCCAACTTTCGC
CTGGAGCTCGTCAACCATCACAACAAATGCAAATCGCATCGCATGTGCTACG
AGCGTGAGCTCCAGGAGCAGCAGCAACATCCCATGCCCAGTGGAAGTGCAG
AGGCAGCTCCAAGCGACGTTCCCCCGAAATCATCACCATAAATGTGGGCAAG
AACAGCGCATAGAGACAGGTCCTTGTAACAGACCCAACACACCGCTCAGGAT
TAGCCATAGGCAATAGCATTAAGTTTTCGTTAACAATAAATGTCCGCAAAGT
AAAATATATTTAGAAACCTAAGAACTTGGCAGCTACTGTGAACAAAAACAAA
CATCATTGACAGCATTTGCATTGGCAAAAATTGAACAATAAACTGAGGAGCA
TTAA

Appendix B: Egg plate recipe

Recipe for grape or apple juice agar plates used in the fly cages to collect eggs, obtained from the lab of Dr. Percival-Smith.

Ingredients	2 litres (~ 50 egg plates or 2 packages)
Water	1000 mL
100% apple or grape juice	1000 mL
Agar	50 g
Sugar	120 g
Propionic acid	6 mL

Add water, juice, and agar together in a large pot and put on high heat. Next, add agar and stir constantly as agar burns quickly, followed by propionic acid (wear gloves). Heat to a boil and keep stirring until no granules are left and most of the bubbles have dissipated. Turn the heat down and keep stirring until the solution appears less opaque. Distribute into 100 mL petri dishes and let it cool and harden before placing in fridge.

Appendix C: Reciprocal hemizygosity test

D. melanogaster and *D. simulans* flies with *fru* disrupted (maintained over either a balancer –*TM6B* – or inversion) will be crossed to the original stocks used for embryo injections of the opposite species (Figure 8). The two types of interspecies female hybrids that are produced [*mel(fru-)/sim* and *mel/sim(fru-)*] will be assayed. Pure species hybrids (*mel/sim*) made from the same stocks will be assayed at the same time as a positive control. Mating assays will follow that of Laturney and Moehring (2012b). Briefly, courtship and mating will be observed for one hour at 24°C with 70% relative humidity, between virgin female hybrids and virgin wild-type *D. melanogaster* males with GFP-tagged sperm in 30 mL glass vials. After one hour, the pairs will be placed into 30 mL food vials for 24 hours. After 24 hours, the female reproductive tract will be dissected to score for presence of sperm (as a proxy for mating) or absence of sperm (as a proxy for no mating) to determine if there is a change in behavioural phenotype to reject a mate. Proportion of pairs that mated will be analyzed using a z-test. If *D. simulans fru* increases female rejection of *D. melanogaster* males, as expected, then the unmasking of this allele in *mel(fru-)/sim* hybrids should result in significantly reduced mating compared to *mel/sim(fru-)* hybrids and *mel/sim* hybrids (Figure 8).

Curriculum Vitae

Name: Jalina Bielaska Da Silva

Post-secondary   University of Western Ontario
Education and   London, Ontario, Canada
Degrees: 2011-2015 Honours B.Sc. with distinction

University of Western Ontario
London, Ontario, Canada
2015-2017 M.Sc.

Honours and   Certificate in University Teaching and Learning
Awards   University of Western Ontario
2017

Fall Travel Award,
University of Western Ontario
2016

Best Graduate Lightning Talk
OE3C, University of Toronto
2016

Outstanding Undergraduate Thesis Presentation
Biology Day, University of Western Ontario
2015

University of Western Ontario Scholarship of Distinction
2015-2016

Related Work   Teaching Assistant:
Experience   BIO 3467 Evolution and Reproduction, BIO 3598 Behavioural
Genetics, BIO 2601Organismal Physiology, BIO 2581 Genetics
The University of Western Ontario
2015-2017

Presentations:
Bielaska Da Silva, J. & Moehring, A. (2017). Characterizing the genetic basis of
behavioural isolation in *Drosophila simulans* using CRISPR/Cas9. 58th Annual
Drosophila Research Conference, San Diego, CA, USA. [Poster].
Bielaska Da Silva, J. & Moehring, A. (2016). Characterizing the genetic basis of
behavioural isolation in *Drosophila simulans* using CRISPR/Cas9. Ontario Ecology,
Ethology, and Evolutionary Conference, Toronto, ON, CA. [Oral presentation].