Post mating isolating barriers between Drosophila species and the role of seminal fluid gene expression

By

Carolina Flacchi

…..

MASTER OF SCIENCE

Department of Biology

University of Winnipeg

Winnipeg, Manitoba, Canada

Abstract

Reproductive genes are known to be among the fastest evolving category of genes within the genome, and males' reproductive genes show a high divergence between species. A class of genes expressed in the male's reproductive tract, the Seminal Fluid Protein genes (SFPs), have been shown to be the most rapidly evolving male genes. The fast evolution and divergence of these genes were first attributed to forms of postcopulatory sexual selection and sexual conflict. However, a recent study that analyzed the molecular evolution of SFPs among different species of *Drosophila* at the coding sequence level found that the responsible force driving the rapid evolution of SFPs was relaxed selection, with only a small proportion evolving by positive selection. In this thesis, I focus on analyzing the molecular evolution of SFP genes at the gene expression level, rather than at the coding sequence level, and on understanding whether changes at the gene expression level can trigger the evolution of reproductive barriers. From my analysis, it emerged that SFP genes show a higher divergence in expression compared to the average accessory gland gene, but do not show a higher polymorphism. Moreover, after knocking down four genes under positive selection, that have been previously shown to affect intraspecific sperm competition, and after performing mating experiments both at the intra and interspecific level between *D. melanogaster* and *D. simulans*, we find no difference in the ability of the *D. melanogaster* knockdowns to outcompete heterospecific *D. simulans* males. These results suggest that genes influencing forms of postcopulatory sexual selection (i.e., sperm competition) and those influencing reproductive barriers (i.e., conspecific sperm precedence) do not share a common genetic basis. This hints to the possibility that, while intraspecific sperm competition can be

affected by the perturbation of a single gene, the perturbation of more than one gene at the same time might be needed to affect conspecific sperm precedence.

Acknowledgments

I would like to thank my supervisor, Dr. Alberto Civetta, for his guidance throughout the course of my degree. His patience and support were essential for the completion of this work. His belief in my capabilities pushed me in achieving my goals.

Thank you to my committee members, Dr. Jens Franck and Dr. Sara Good for providing helpful suggestions regarding my project.

I want to thank my family for their encouragement and for always being there to support me in all the decisions I made. Thank you for always being proud of my achievements. I couldn't have done anything of this without you by my side.

I am thankful for Erick, you always made me believe in myself, even when I was doubting of my own capabilities. I always found comfort in your words and presence, whenever I needed it.

A special thank you goes to my friends, Anouska and Avinash. These two years wouldn't have been the same without you. Thank you for all the good times, and for standing next to me throughout the bad ones.

Table of Contents

List of abbreviations

SFP: Seminal Fluid Protein

PMPZ: Post Mating Pre Zygotic

CSP: Conspecific Sperm Precedence

PCSS: Post Copulatory Sexual Selection

GFP: Green Fluorescent Protein

RFP: Red Fluorescent Protein

SR: Seminal Receptacle

ST: Spermathecae

SP: Sex Peptide

QTL: Quantitative Trait Loci

ACP: Accessory Gland Protein

SVS2: Seminal Vesicle Protein 2

DGRP: *Drosophila* Genetic Reference Panel

TPM: Transcript Per Million

MSBW: Mean Sum Of Squares Between Species

MSW: Mean Sum Of Squares Within Species

KD: Knockdown

Sb: Stubble

IDT: Integrated DNA Technologies

P1: Sperm Defense

P2: Sperm Offense

ISC: Intraspecific Sperm Competition

FET: Fischer Exact Test

NA: Not Available

GMO: Genetically Modified Organisms

List of Figures

List of tables

1.0 Introduction

1.1 Speciation and Reproductive Isolation

One of the primary leading evolutionary biologists of the 20th century, Ernst Mayr, defined species as "groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups" (Mayr, 1942), establishing the "biological species concept". Another important evolutionary biologist of the same period, Theodosius Dobzhansky, stated that "the biological species is the largest and most inclusive Mendelian population", where a "Mendelian population is a reproductive community of sexual cross-fertilizing individuals which share in a common gene pool" (Dobzhansky, 1950). Today, most people associate the term "species" to "the basic category of biological classification" (Flexner and Hauck, 1993). The definition of species presented by Mayr, commonly referred to as the biological species concept, has, throughout the years, gone through different critiques, and alternative definitions have been proposed (Simpson, 1951; Simpson, 1961). For the purpose of sexually reproducing organisms, the biological species concept serves as a useful definition of species, that emphasizes mechanisms that prevent gene flow between organisms as the key isolating factor. Given the central role of species in systematics and evolutionary biology, identifying the driving mechanisms of speciation is essential to understand biodiversity.

A major interest in speciation studies is directed towards identifying reproductive isolation barriers. Once identified, several questions follow such as understanding mechanisms involved in the establishment of reproductive isolation, deciphering patterns and mode of evolution of isolation mechanisms, and teasing apart the genetic basis behind such mechanisms. For example, Dobzhansky stated that an in depth understanding of the mechanisms of speciation could not be

reached without fully comprehending the genetics of reproductive isolating barriers (Dobzhansky, 1937).

Traditionally, mechanisms of speciation have been classified based on geographic relationship between diverging populations and nascent species (Coyne and Orr 2004). Allopatric speciation, in which geographical barriers cause a large population to split causing the genetic exchange between them to stop, is considered the most common geographic mechanism promoting divergence and speciation (Coyne and Orr, 2004). Under allopatric speciation, several genetic differences between the two, now physically isolated, populations will accumulate leading to the evolution of reproductive incompatibilities between them. These incompatibilities will eventually cause biological incompatibilities upon contact between individuals from different locations (Palumbi, 1994). Understanding what biological changes contribute to the evolution of reproductive isolation is essential to fully comprehend the origin of the species.

1.2 Biological modes of reproductive isolation: pre-mating and post-mating barriers

Genetic differences that accumulate between populations contribute to changes in reproductive traits, and they can have profound implications in the establishment of biological barriers to reproduction, which will, eventually, lead to the establishment of reproductive isolation. Evolutionary biologists have broadly classified biological mechanisms of reproductive isolation into pre-mating and post-mating reproductive barriers (Coyne and Orr, 2004).

Pre-mating reproductive barrier occur before copulation, and they can arise due to ecological, behavioral and even anatomical incompatibilities. An interesting and curious example of a pre-mating reproductive barrier, with behavioral basis, is found in birds, where songs are used

to choose and attract individuals to mate with. A study conducted on passerine birds (Freeman *et al*., 2022) revealed that in two closely related geographically isolated populations (allopatric populations), birds responded strongly to their local song, while completely ignoring the allopatric one, confirming the hypothesis that song in birds represents a powerful pre-mating reproductive barrier. Another example comes from mimicry in butterflies, a trait of critical importance in survival that is also very important in mating behaviour. The study found that closely related *Heliconius* species, which show different mimicry patterns, tend to approach and mate with individuals that express the same colour pattern, rather than choosing an individual with a different pattern of mimicry (Jiggins *et al.*, 2001).

Post-mating reproductive barriers happen after copulation and can additionally be differentiated into two other categories: post-mating prezygotic reproductive barriers, and post-mating postzygotic reproductive barriers (Coyne and Orr, 2004). Post-mating pre-zygotic (PMPZ) barriers happen after copulation, but before a zygote is formed. The complex interactions between the reproductive traits of males and females can lead to several difficulties and impediments, especially in heterospecific crosses, and they are major factors in the establishment of PMPZ reproductive barriers. A widely known form of post-mating prezygotic barrier is conspecific sperm precedence (CSP). This reproductive barrier is particularly noticeable in polyandrous species such as insects and birds, where females mate with more than one male, which creates opportunities for competition between sperm and for female cryptic choice of sperm for fertilization. CSP occurs when a female successfully mates with both a conspecific and a heterospecific male, but most of the progeny is fathered by the conspecific male, regardless of the mating order (Price, 1997). This creates a barrier to gene flow between related species that are still not completely reproductively isolated from each other (Price, 1997; Price *et al.*, 2000). Conspecific sperm precedence has been

observed in a wide variety of species, both terrestrial (Price, 1997; Robinson *et al*., 1994; Howard *et al*., 2002) and marine (Geyer and Palumbi, 2005), and it has also been detected in plant species (Cruzan and Arnold, 1993), where is known as conspecific pollen precedence. When heterospecific pollen and conspecific pollen are found in the same stigmas, hybrid seeds are produced in much lower quantities (Howard, 1999).

Post-mating post-zygotic barriers happen after copulation and after the zygote formation. These barriers usually result in hybrid inviability or in the production of viable but sterile offspring. One of the most well-known examples of hybrid sterility is the mule, the F1 progeny from the mating of a horse and a donkey, which is marked by a deficiency of males, all sterile (Short, 1975).

Studies that have compared the amount of genetic divergence with the origin of different modes of isolation have established that, on average, pre-mating barriers evolve faster than post-mating ones (Coyne and Orr, 1989; Turissini *et al*., 2018). However, this does not exclude post-mating barriers from having an important role in speciation. For example, a study conducted on three distinct populations of *Drosophila montana* revealed that premating barriers were stronger among geographically distant populations. However, PMPZ barriers showed no association with distance, suggesting a critical role in preventing hybridization among individuals from nearby distinct populations (Garlovski and Snook, 2018). Another example supporting an important role of post-mating barriers in early stages of speciation comes from a study performed on two *Drosophila* subspecies, *D. w. willistoni* (North America, Central America, and northern Caribbean islands) and *D.w. winge* (South America and southern Caribbean islands). The study detected significant levels of premating isolation between geographically distant populations, but not between geographically closer populations of the same subspecies. This result shows that premating barriers can have a fast rate of evolution between species. However, it also suggests a

more prevalent role of post-mating isolation in the early stages of speciation. When the two subspecies are crossed, the male offspring derived from one direction of the crosses is sterile. These sterile hybrids, when mating, manage to transfer the ejaculate to the female's reproductive tract, causing an elongation and expansion of the uterus. However, due to a morphological atrophy at the basal end of the testes, their ejaculate lacks sperm, as it is not properly transferred to the sperm pump (Davis *et al*., 2020).

Natural selection can, by favouring different adaptations to different environments, lead to reproductive isolation between diverging populations (Schluter, 2000; Coyne and Orr 2004; Nosil, 2012). However, another force that has been suggested having an important influence on the establishment of reproductive isolation is sexual selection. Sexual selection would lead to reproductive isolation by causing significant divergence in reproductive traits, such as male reproductive traits contributing to male success and female preference (Lande, 1981; Maan and Seehausen, 2011). Because of its role in driving phenotypic diversification and increasing genetic divergence, some studies are highly supportive of the idea of sexual selection being a powerful driver of speciation (Panhuis *et al.*, 2001; Boughman, 2001; Mendelson and Safran 2021). However, because of the lack of proof of sexual selection directly driving reproductive isolation, others have argued that it is unlikely for sexual selection to act on its own and have rather suggested that it must act as a secondary force alongside other selective forces as, for example, natural and ecological selection (Maan and Seehausen, 2011; Langerhans *et al.*, 2013; Safran *et al.*, 2013).

It is for this reason that finding commonalities between mechanisms of sexual selection and reproductive isolation, as well as common genetics underpinning, would provide grounds for a link between sexual selection and speciation.

1.3 Mechanisms of sperm competition

The idea of sexual selection being a process able to exclusively act before copulation remained solid until Geoff Parker's work on insects suggested that sexual selection could continue after copulation (Parker, 1970). Parker observed a pattern of polyandry across different species of insects. Polyandry is expected to favour male – male competition and female choice among sperm in the storage organs of the females after copulation and even during fertilization. In fact, females can store sperm from different males for several days after mating (Neubaum and Wolfner, 1999; Neubaum and Wolfner, 1999; Perotti, 1973). This characteristic sets the stage for a form of post copulatory sexual selection (PCSS) known as sperm competition. Sperm competition was defined by Parker as "the competition within a single female between the sperm from two or more males for the fertilization of the ova" (Parker, G. 1970).

PCSS includes selection *via* sperm competition and cryptic female choice, which refers to the ability of the female to bias sperm from different males to be used for fertilization (Thornhill,1983). Since Parker's discoveries, sperm competition has been shown to occur in a wide variety of species. It has been observed in mammals, such as in a species of sheep (Preston *et al*., 2003), squirrels (Boellstroff *et al*., 1994) and mice (Firman and Simmons, 2008) and it also proved to occur in birds (Birkhead, 1998), insects (Parker, 1970; Simmons, 2001) and nematodes (LaMunyon and Ward, 1999).

Parker's work on different classes of insects showed that in most, the last male to mate fathers most of the offspring, with a fertilization success of 75%-100% compared to the male that mated first. This phenomenon is known as last male sperm precedence (Parker, 1970). The fertilization success of a male that has mated last to a female suggests the involvement of different

possible mechanisms, such as sperm stratification favouring the last male sperm in fertilization, a mechanism often referred to as "first in last out" or last male sperm displacement of stored sperm. While mechanisms can be inferred from paternity scores in competitive settings (Simmons, 2001; Simmons, 2002), the use of transgenic *Drosophila melanogaster* that expressed green-fluorescent sperm has facilitated the resolution of mechanisms of sperm competition. When *Drosophila* females were firstly crossed to a GFP-male and then crossed to a second male, which did not express GFP-sperm, the amount of fluorescent sperm in her storage organs was much lower when compared to females who were single-mated to GFP-males, suggesting that sperm from the second male had the ability to displace rival stored sperm (Civetta, 1999). After copulation, the transferred ejaculate follows a specific pathway within the female's reproductive tract. Sperm are deposited into the uterus (bursa), before entering the storage organs (a seminal receptacle - SR and two spermathecae - ST), which will make them accessible for fertilization (Nonidez, 1920). A study has allowed for the visualization of sperm competition *in vivo*, by using both GFP and RFP-transgenic (red fluorescent protein) lines of *Drosophila,* which could be easily differentiated within the female's reproductive tract (Manier *et al.*, 2010). Their results showed the involvement of two mechanisms by which resident sperm was being displaced. The first one was a female-mediated mechanism that did not involve second male's sperm, while the second one did. The female mediated mechanism of sperm ejection confirmed the active role of females in sperm competition that had been previously inferred from work that competed the same two male genotypes using different female genotypes and observed extensive variation in paternity success depending on the female's genotype (Clark *et al*., 1999). The second mechanism confirmed competition between the different sperm sources. They observed that as second male's sperm entered the female's reproductive tract, resident sperm was being displaced from the SR and the ST back to the bursa,

allowing for the second male's sperm to be the first one used in fertilization and, therefore, for last male precedence (Manier *et al*., 2010).

Mechanistically, we have some understanding of how males manage to displace sperm from rivals. The displacement of a stored sperm is often due to a physical manipulation of the resident sperm by the second male. Sperm displacement can be achieved by the influence of specific components of the seminal fluid (i.e., seminal proteins). Seminal fluid can have major effects on resident sperm, and it can even affect rival sperm motility (Liberti *et al*., 2018). A study conducted on a species of polyandrous ant (*A. echinatior*) compared sperm motility parameters between sperm who were exposed only to own seminal fluid and sperm exposed to mixed seminal fluid (both own and rival). Their results showed an increase by 50% in sperm motility in mixed ejaculates, suggesting that enhancing sperm motility is costly to males (Liberti *et al*., 2018). Early studies in *Drosophila* have suggested that the second male seminal fluids could incapacitate and even kill resident sperm (Harshaman and Prout, 1994; Price *et al.*, 2000). The idea of sperm incapacitation has been more recently disproved and there has been evidence of a beneficial role of seminal fluid proteins on rival sperm (Holman, 2009; Simmons, 2011; Misra and Wolfner, 2020). However, in ants (*Atta colombica*), sperm viability was shown to be affected by rival seminal fluid (Dosselli *et al.*, 2019).

Specialized structures on the males' genitalia confer them the ability to physically remove rival sperm form the female's storage organs, giving them the chance to replace it with their own (Waage, 1986). The mechanism by which second male's sperm replaces rival sperm from the storage organs of the female is known as sperm offense, while the one by which a stored sperm resists this displacement is known as sperm defense. Both the processes of sperm offense and sperm defense act as selective pressures on males which lead to the evolution of new reproductive

traits and mechanisms that can increase the male's reproductive fitness. These mechanisms can be grouped based on their goals. The first one is to promote sperm success in direct competition, while the second serves to prevent competition from happening. An important adaptive mechanism that prevents competition is known as "mate guarding". Mate guarding refers to a situation where copulation continues even after insemination has been completed, preventing females to remate with any other male until all the eggs have been laid (Grafen and Ridley, 1983; Alcock, 1994). In some species of insects, copulation was observed to last up to 11 days (Carroll, 1991). Mate guarding can persist even when the male is not physically present. This happens when anti-aphrodisiac pheromones are transferred to the female, which makes her unattractive to other males (Mozuraitis *et al*., 2019). A similar process is the formation of a mating plug, which is the result of the coagulated male's ejaculate that blocks the female's reproductive tract and prevents sperm from rival males to enter the sperm storage organs (Polak *et al.*, 2001). Seminal fluid's components (i.e., seminal fluid proteins) can also have direct effects in lowering the chances of competition between rival sperm. An example is the Drosophila melanogaster *sex-peptide* (SP) seminal fluid protein, which is known to have a role in lowering female's receptivity to remate. Studies have demonstrated that females which mated with knock-down males for *SP* had a higher receptivity to remate, allowing for a second copulation with a different male to happen at a faster rate (Chapman *et al.*, 2003).

Parker predicted that, at each copulation, the amount of sperm produced by males will depend on the level of sperm competition risk they faced (Parker, 1990). In direct competition, when the risk of competition is high, males can produce a larger amount of sperm to increase their reproductive success (Wigby and Chapman, 2004). A previously mentioned study (Manier *et al*., 2010) also reported that, in competitive settings, the ejaculate size of the second male was highly

correlated to the amount of resident stored sperm that was being displaced. Moreover, a study conducted on different species of *Cephalopods* (i.e., squid) showed that, in some species, an increased risk of sperm competition resulted in a greater sperm allocation in the female's storage organs (Iwata *et al*., 2021).

Theory also predicts that with increased levels of sperm competition risk, an increased testis investment is also expected (Parker, 1998). This prediction has already been confirmed by studies on both vertebrates and invertebrates (Rowley *et al*., 2019; Simmons and García-González, 2008), and further research has supported the hypothesis that a greater testis investment increases a male's reproductive success (Satoshi *et al.,* 2006). A greater testis investment would allow to produce bigger or more sperm, which could have higher chances in sperm competitive settings. In *Drosophila*, males with larger accessory glands transferred a greater amount of *sex-peptide*, suggesting that, in competitive settings, variation in the production of specific seminal fluid proteins could also have essential roles in increasing reproductive success in males (Wigby *et al*., 2009).

1.4 The genetic basis of sperm competition

The use of paternity assay experiments has contributed to the understanding of the phenotypic effects of sperm competition. However, the genetic basis underlying this process are still to be fully uncovered. The identification of genes and molecular pathways associated with sperm competition has been a major challenge due to the polygenic nature of the process, the complex genetic variance inheritance, and the lack of availability of transcriptomic and genomic data (Civetta and Ranz, 2019). Elucidating the genetic processes underlying sperm competition would identify specific genes that play major roles in this form of sexual selection, and therefore,

in driving the evolution of sexual traits. In addition, expanding knowledge on these genes could serve as a tool to test whether sperm competition plays a role in speciation.

The different approaches that have attempted to identify the genetic basis of sperm competition have been recently reviewed (Civetta and Ranz, 2019). Quantitative Trait Loci (QTL) mapping studies have failed in identifying specific genes with major roles in sperm competition, suggesting, instead, the involvement of multiple genes and genetic factors (Hughes, 1997; Lawniczak and Begun, 2005; Hughes and Leips, 2006). A QTL mapping study conducted in different species of *Peromyscus* mice has, however, managed to identify a specific gene of large effect, the *protein kinase cAMP-dependent regulatory type I alpha* or *Prkar1a* (Fisher *et al*., 2016). The gene was shown to be correlated with the phenotypic difference in the sperm midpiece length between species which influences sperm swimming velocity. Their results suggested that allelic variation at this locus can impact male's fitness (Fisher *et al*., 2016).

Another approach is to conduct gene-phenotype association studies that are used to identify genetic variants (i.e., single-nucleotide polymorphisms) at candidate genes, with known functions in non-competitive reproductive functions, that associate with differences in sperm competitive abilities. Some of these candidates include those that are known to have functions in sperm storage, sperm release or in regulating receptivity in females. Gene-phenotype association studies identified a role for ACPs (accessory gland proteins) alleles in sperm competition (Clark *et al*., 1995; Fiumera *et al*., 2005; Fiumera *et al*., 2007; Wong *et al*., 2008; Greenspan and Clark, 2011; Zhang *et al*., 2013). However, this approach suffers from several limitations. The results could be different depending on population-specific genomic backgrounds, and the level of stringency used might produce different associations and even false positives. Ultimately, the results of association studies should be validated by follow-up functional assays, which include gene-perturbation

experiments such as gene knockouts, knockdowns and gene editing. The use of these type of gene manipulation assays has highlighted important role of SFPs in reproduction.

1.5 Functions of seminal fluid proteins (SFPs)

A particular class of reproductive genes are those contained in the seminal fluid of the males, which encode a group of proteins that are transferred to the females after copulation, known as seminal fluid proteins (SFPs). SFPs are produced in the male's reproductive tract, mostly by the secretory cells of the accessory glands (AGs), and it is now known that, in *Drosophila*, the seminal fluid is enriched with as many as 300 different seminal fluid proteins (Findlay *et al*,. 2008; Ravi Ram and Wolfner 2007; Wigby *et al.*, 2020; Hurtado *et al.*, 2022). Males transfer both sperm and seminal fluids during copulation, and the seminal fluid triggers a series of physiological response in females (Wolfner, 1997). The seminal fluid is known to be a cocktail of different substances, such as water, lipids, proteins, carbohydrates, hormones and microbes (Hopkins *et al.*, 2017).

Drosophila melanogaster has been a powerful genetic tool to study the effects of SFPs in females after mating. Many SFP functions are still unknown, however, molecular and genetic techniques that allowed for the identification and characterization of some of them, have proved the essential role of these proteins in pre- and post-mating fitness. One of the first seminal fluid proteins to be identified is known as *SP* (Chen and Bühler, 1970). *SP* binds to sperm when the seminal fluid is transferred to the female's reproductive tract, which allows *SP* to stay in the female's reproductive tract for a continued period (Saudan *et al.*, 2002; Peng *et al.*, 2005). Normally, females that mate with normal males become unreceptive for 5 days after the first mating (Manning, 1962). However, females that mate with spermless males become reluctant to remate only for 1-2 days (Manning, 1967). This suggests that the long-term response of *SP* is

provided by its binding to the sperm, and absence of this bound only generates a short-term response. More recent studies have identified that the binding between *SP* and sperm is accomplished by the interaction of several different proteins that form the "Sex-peptide network" (Ravi Ram *et a.*, 2009; Findlay *et al.*, 2014). *CG1652* and *CG1656* (C-type lectins), *Sems* and *CG9997* (serine proteases) and two cysteine rich secretory proteins (*CG17575*, *Antares*) were found to be part of the sex peptide network and observed to be binding to sperm in the female's reproductive tract within the first two hours after mating (Findlay *et al.*, 2014; LaFlamme *et al.*, 2012; Ravi Ram and Wolfner, 2009). These SFPs work together to facilitate the binding of *sex-peptide* to the sperm, and the absence of any of these proteins results in an incorrect retainment of *sex-peptide* in the female's reproductive tract, which, in return, does not trigger the post-mating responses it is involved in (Findlay *et al.*, 2014; Peng *et al.*, 2005; Pilpel *et al.*, 2008; Ravi Ram and Wolfner, 2007, 2009).

One of the most important processes seminal fluid proteins are involved in is, however, sperm competition. Shortly after mating, the sperm transferred to the female is stored in her storage organs, where it can stay for a long period of time (Manning, 1962; Civetta, 1999). Sperm storage therefore provides the grounds for the competition between sperm from different males, within the female's reproductive tract. Acps were identified as essential factors in sperm storage mechanisms (Tram and Wolfner, 1999). An example of an accessory gland protein with a crucial function in sperm storage is *Acp36DE* (Neubaum & Wolfner 1999). *Acp36DE* is found in the anterior part of the mating plug, and it helps the sperm move close to the storage organs by tightly associating with the sperm mass and by entering the sperm storage organs (Bertram *et al.*, 1996; Neubaum & Wolfner 1999; Avila and Wolfner, 2009). Studies have demonstrated the involvement of *Acp36DE* in sperm competition by using *Acp36DE*-deficient males in competitive mating experiments

(Chapman *et al.*, 2000). By using *Acp36DE* mutant males as second to mate, it was shown that the alteration of sperm storage affects the outcome of sperm competition. In fact, null males were showing much lower P_2 (proportion of progeny fathered by the second male to mate) values compared to control males, suggesting an involvement of *Acp36DE* in rival sperm displacement (Chapman *et al.*, 2000). Moreover, another study has confirmed the role of *Acp36DE* in sperm competition in heterospecific crosses, suggesting its role in reproductive isolating barriers (Castillo and Moyle, 2014). Their results reported that *Acp36DE*-deficient males, when used as second males in heterospecific crosses, would achieve much lower amount of progeny compared to control males. The same study has, in addition, investigated the roles in sperm competition of two other accessory gland proteins, *sex-peptide* and *CG9997* (Castillo and Moyle, 2014). Apart from being an essential player in the regulation of egg production and female receptivity, *sex-peptide* also functions in the mechanism of sperm release from female storage (Avila *et al.*, 2010). On the other hand, *CG9997* is involved in the sex peptide network where it transfers three other Acps required for *sex-peptide* to bind to sperm, and studies have demonstrated that a perturbation of this gene affects the sperm release from the female storage organs (Ram and Wolfner, 2007). By using sperm competition assays in both heterospecific (conspecific sperm precedence) and conspecific crossings (intraspecific sperm competition), the study confirmed an offensive role in sperm competition both for *sex-peptide* and *CG9997*. Moreover, *CG9997* resulted to have a phenotypic effect on CSP. In heterospecific crosses where null mutant males for *CG9997* were crossed as second males, the proportion of the first heterospecific male's progeny would increase, suggesting a common genetic basis between mechanisms of sperm competition and reproductive barriers (Castillo and Moyle, 2014).

A recent review has identified 33 genes in flies affecting sperm competition based on studies that have used a combination of approaches (Civetta and Ranz, 2019). Among the 33 genes identified, 10 were genes expressed in the male accessory glands that produce the seminal fluid. Moreover, the review also identifies genes with roles in sperm competition in other species. For example, the amount of protein produced by the *seminal vesicle protein 2* (*SVS2*) gene was demonstrated to be positively correlated to the risk of sperm competition in mice (Ramm *et al.*, 2015; Kawano *et al.*, 2014). Furthermore, absence of this gene's product would cause sperm fracture and death (Ramm *et al.*, 2015; Kawano *et al.*, 2014).

1.6 SFPs can modulate a wide variety of post-mating physiological responses in females

The transfer of accessory gland proteins (Acps) to the female's reproductive tract can also have an impact on different aspects of the female's physiology. It was observed that females with a higher mating rate show a much lower lifespan compared to those that didn't mate, or that mated at a lower frequency (Chapman *et al.*, 1995). Specifically, females who mated to males lacking Acps were showing a much lower mortality compared to those who mated to normal males, confirming the hypothesis that these proteins are the main factor in the cost of mating in females (Chapman *et al.*, 1995). Moreover, a survey of males from 51 chromosome-extracted *D. melanogaster* lines found significant differences in longevity of females mated to males of different genotypes. Females' mortality was found to significantly correlate with the proportion of progeny sired by the first male to mate relative to tester males, supporting the hypothesis of a tradeoff between defensive sperm-competitive ability of males and life-history parameters of mated females (Civetta and Clark, 2000). The mechanisms by which Acps cause a decreased longevity in female could be explained by the fact that they can enter the female's circulatory

system. Some Acps have been shown to be able to enter the female's hemolymph shortly after mating (Lung and Wolfner, 1999). Here, these proteins can interfere with biochemical processes which can affect female's physiology (Wolfner, 1997; Monsma *et al.*, 1990).

Other effects that SFPs can have on broader aspects on female's physiology include immunity and abnormal feeding behaviour (Avila *et al.*, 2011). Gene expression studies have demonstrated that the Acps can upregulate the expression of different antimicrobial genes (McGraw *et al.*, 2004). Moreover, three specific Acps (*CG6168*, *CG9334*, and *CG10284*) were proved to help reduce a bacterial infection of infected females after mating (Mueller *et al.*, 2006). Mated females were also observed to feed at higher rates than virgin females (Carvalho *et al*., 2006). Their results revealed an essential involvement of *SP* in this post-mating response. The change in feeding behaviour and stimulation of food intake would not happen in females mated to males who lacked the expression of *SP* (Carvalho *et al.*, 2006).

1.7 The evolution of SFPs

Reproductive genes are known to be one of the fastest evolving categories of genes in the entire genome (Swanson *et al.*, 2001; Swanson and Vacquier, 2002; Haerty *et al.*, 2007). Because of their rapid evolutionary rate, specific classes of reproductive genes are used to try to understand the evolution of reproductive traits (Wigby and Chapman, 2005; Bono *et al.*, 2015). Moreover, the rapid evolution of reproductive genes is interesting as it could contribute to the establishment of reproductive isolation between species (Martin and Hosken, 2003; Orr, 2005; Nakadera *et al.*, 2020; Garlovsky *et al.*, 2020).

The rapid evolutionary rate that characterizes changes in the coding sequence of seminal fluid proteins has been frequently attributed to forms of strong directional, positive selection

(Civetta and Singh 1998; Begun *et al.*, 2000; Holloway and Begun 2004; Haerty *et al.*, 2007). Most studies have tested the evolutionary forces driving the evolution of SFPs by comparing the ratio of nonsynonymous substitutions (d_N) to synonymous ones (d_S) between species or in a phylogenetic context without examination of patterns of polymorphism within species. Analyses that combine information on both polymorphism within species and divergence between species could provide different views on the evolution of these reproductive genes. In fact, a recent review has shown that by incorporating polymorphisms into the analysis, the pattern of rapid evolution of many reproductive genes can be explained by relaxed purifying selection (Dapper and Wade 2020). A more recent study (Patlar *et al.*, 2021) has used a population genetics approach that incorporated polymorphism information from both ancestral and derived populations of *Drosophila melanogaster*, along with sequence data information from its close relative *D. simulans*, to evaluate the mode of evolution of 317 SFPs, most of them known to be transferred to the females after mating (Wigby *et al.*, 2020). By including in their statistical analyses both polymorphism within species and divergence between species data, their results confirmed the faster evolution of most SFP genes compared to the rest of the genome but showed that a significantly high portion of SFP genes evolve rapidly under relaxed selection (50-57%), with only a smaller percentage under positive selection (7-12%) (Patlar *et al.*, 2021).

The finding of relaxed selection being the main force in driving the rapid divergence of SFPs is intriguing, giving their important role in reproductive fitness (Ravi Ram and Wolfner 2007; Avila *et al.*, 2011; Sirot *et al.*, 2014; Schjenken and Robertson 2020; Wigby *et al.*, 2020). This result has led us to speculate (Patlar and Civetta, 2021) that changes at the gene regulation and expression level could be perhaps responsible for adaptive molecular and phenotypic differences in reproductive strategies between species

1.8 Mechanisms of conspecific sperm precedence

Conspecific sperm precedence is, as previously mentioned, a widespread form of post-mating pre-zygotic reproductive barrier (Price, 1997; Price, 2000). However, the molecular and genetic mechanisms underlying this process have remained mostly unknown due to the difficulties in visualizing the female's reproductive tract, where the conspecific sperm is favoured against the heterospecific one. By using differently labeled sperm heads, a study in particular has contributed to our current understanding of the mechanisms of CSP (Manier, *et al.*, 2013). The study found that the amount of sperm transferred by a conspecific male drop significantly when preceded by a heterospecific male. However, when *D. simulans* males and *D. mauritiana* males were crossed with a *D. simulans* female, the *D. simulans* sperm was shown to have a better displacing ability compared to the heterospecific male, even when mated second and despite transferring a reduced number of sperm. Moreover, females were shown to play an active role in favouring conspecific sperm. In crosses where the heterospecific males were mated as second, their ejaculate was being ejected by the females within the first hour after mating. However, when conspecific males were used as second to mate, females would retain their ejaculate for a long time (Manier, *et al.*, 2013). It is also known that storage organs in the female's reproductive tract are first or second-male biased (Manier *et al*., 2013). Specifically, the SR is biased toward the first-male sperm, while the spermathecae has a second-male sperm biased. In addition, in *Drosophila*, the organ responsible for the short-term sperm storage is the SR, while the spermathecae are responsible for the long-term storage (Neubaum and Wolfner, 1999; Pitnick *et al.*, 1999). Females who mated first with an heterospecific males were observed to favour the spermathecae, while in crosses where the

conspecific male was first, females favoured the SR. In both these situations, females were actively favouring the conspecific sperm for the fertilization success (Manier, *et al.*, 2013).

All these highlighted mechanisms suggest an interaction between males and females' reproductive traits, hinting at a strong possibility of post-copulatory sexual selection being involved in the establishment of reproductive isolation and speciation.

2.0 Objectives and hypotheses

The rapid coding sequence evolution of SFP genes has been recently attributed to relaxed selection (Patlar *et al.*, 2021). However, evolutionary patterns of SFP genes at their gene expression level have not been investigated yet, and the rapid evolution of this class of genes could depend on changes in gene regulation and expression. The two main objectives of my thesis are to test the mode of evolution of SFP gene expression and to test whether experimental manipulation of positively selected SFP genes can break down barriers of interspecies post-mating prezygotic isolation.

The first approach of my thesis was to identify genes that showed major differences in gene expression between the two closely related species *D. melanogaster* and *D. simulans*. Specifically, I tried to identify genes that lacked orthologs in *D. simulans*, to perform a genetic manipulation of these in *D. melanogaster* and observe any differences in the P1 and P2 ability of the knockdowns in heterospecific settings (CSP).

The second approach was to use published male-accessory gland transcriptomic data obtained from different strains of *Drosophila melanogaster* and its close relative *Drosophila simulans* (Cridland *et al.*, 2020), to identify the mode of evolution of accessory gland expressed genes. I hypothesized that, given their important role in reproduction, most SFP genes should show either rapid divergence between species driven by positive directional selection or conserved levels of expression between species maintained by negative purifying selection. Previous studies have proposed that postcopulatory sexual selection could drive the rapid evolution of SFP genes leading to the establishment of post-mating prezygotic barriers to reproductive isolation between species. Genes under positive directional selection fueling differences in expression between species, and

those with evidence for a role in sperm competition, a form of post-copulatory sexual selection, are likely candidates to test for effects of gene expression manipulations on post-mating reproductive isolation. Therefore, I focused my gene perturbation assays on specific genes under positive selection that have been previously shown to have roles in intraspecific sperm competition (Civetta and Ranz, 2019; Patlar and Civetta, 2022) and performed sperm competition assays at the interspecific level.

3.0 Methods

3.1 *In silico* search for *D. melanogaster* SFP genes without *D. simulans* orthologs

In order to identify, from a list of 292 SFPs known to be transferred to females during mating (Wigby *et al.*, 2020), genes with the largest difference in the amount of gene expression between two closely related species of *Drosophila*, *D. melanogaster* and *D. simulans*, I searched for those that lacked orthologs (i.e., genes that evolved from a common ancestral gene by speciation) in *D. simulans*. The Batch Download tool available on Flybase (https://flybase.org), a database of Drosophila genes and genomes, was used to query the 292 SFPs for orthologs in other species of *Drosophila*. The Batch Download tool retrieves gene orthologs across *Drosophila* species via the Orthodb database (www.orthodb.org/v9.1/) a catalog of orthologs that uses sequence homology and gene functional annotations to recover gene orthologies. SFP genes that were not annotated in the *D. simulans* species were selected as candidates.

3.2 Experimental test of *D. simulans* lack of orthology

To test the actual absence of expression in *D. simulans* of the genes identified as lacking orthologs in this species, a PCR was performed. For each gene, the FASTA transcript sequence of *D. melanogaster* was downloaded from Flybase, and the BLAST tool within GenBank was used to identify similar *D. simulans* sequences. The Primer3Plus tool (https://www.primer3plus.com) was used to design primers that could amplify "genes" with sequence similarities between the two species. The selected primers were ordered from Integrated DNA Technologies (https://www.idtdna.com/pages), diluted with distilled water upon arrival into 100 nmol stocks and

kept at -20°C. Before PCR reactions were performed, the primers were diluted into 10 nmol working stocks.

To test the target-specificity of each primer for each gene sequence, a PCR using samples of DNA extracted from both *D. melanogaster* and *D. simulans* was performed. DNA was extracted from 10 flies of each species (5 males and 5 females) following a standard lab protocol (**Supplementary Protocol 1.0 and 2.0**). In the PCR reaction, the templates were denatured for 5 min at 95° C. This step was followed by the primer annealing step, where primers were kept at 57 °C for 30 s. The extension was performed at 72° C for 45 s and 34 cycles of amplification were repeated. PCR products were later visualized on a gel imaging machine (iBright Imaging System – ThermoFisher) after running an electrophoresis on a 1.5% agarose gel. Stained nucleic acid bands on the gel could be visualized due to the SYBR Green stain utilized during the PCR reaction (DreamTaq Hot start green PCR MIX, K9021). A 100 bp ladder (GeneDireX, Cat. #DM001-R500) was used as a reference capture the DNA band's size.

Once the primers were validated, RNA was extracted from the reproductive tract of both species to test whether the SFP genes predicted to be unique to *D. melanogaster* truly lacked expression in the reproductive tract of *D. simulans* males. The reproductive tract of 10 adult males (4-5 days old) from each species was dissected, under a dissecting microscope using fine pins and forceps (Fine Science Tools) in a 1X PBS drop (ThermoFisher Scientific, Cat. #SH30256.01). The tissue was then placed in a lysis solution and RNA was extracted using the Aurum Total RNA mini kit (BioRad, Cat. #7326820) following the manufacturer protocol **(Supplementary Protocol 3.0).** The RNA concentration and purity were then checked under the Nanodrop by examining the ratio of absorbance at 230, 260 nm and 280 nm. The 230/260 ratio indicates how pure the sample

is from salts, while the 260/280 indicates any protein contamination. The optimal values for these ratios are respectively 2 for the 260/280 ratio (for RNA samples) and 2 to 2.2 for the 230/260 ratio. RNA extractions were performed on a Winnipeg (Wpg02) strain of *D. melanogaster* and on a wild-type strain (DSSC 14021-0251.269) of *D. simulans*, purchased from the National Drosophila Species Stock Centre (https://www.drosophilaspecies.com/). The extracted RNA was then stored at -80°C. cDNA was synthetized using reverse transcriptase, included in the cDNA synthesis kit (BioRad iScript Select cDNA synthesis, Cat. #1708897) following the manufacturer protocol (**Supplementary Protocol 4.0**). A PCR was performed, following the protocol described above. I tested for presence of transcripts (gene expression) in both *D. melanogaster* and *D. simulans* by gel electrophoresis and staining as described above.

3.3 SFP genes with differential expression between *D. melanogaster* and *D. simulans*

A transcriptomic analysis was performed to identify genes with large changes in the amount of expression between the two species. Available transcriptomic data (Cridland *et al.*, 2020) obtained from the accessory gland of 6 *D. melanogaster* DGRP lines (Raleigh: SRR10253134, SRR10253145, SRR10253156, SRR10253167, SRR10253178 and SRR10253179) and 1 *D. simulans* strain (LARA: SRR10253130). *D. melanogaster* and *D. simulans* sequence reads were uploaded onto Galaxy (galaxy.org) using the "*Download and Extract Reads in FASTA/Q format from NCBI SRA*" tool Version 2.11.0+galaxy0), along with the reference genomes for both species. The *D. melanogaster* (release version r6.42) and the *D. simulans* (release version 2.02) reference genome annotation files (FASTA files) were downloaded from the FlyBase repository (http://ftp.flybase.net/genomes/).

To align the sequence reads of the transcriptomic data to the reference genomes, the HiSat2 tool (Kim *et al.*, 2015) was used with default parameters. In order to assemble the RNA-seq alignments into potential transcripts, the StringTie tool (Pertea *et al.*, 2015), with the "*reference transcript only*" option, was used and the normalized estimates of expression for each gene were obtained as transcript per million (TPM). A list of 317 accessory gland expressed genes that were previously analyzed for patterns of evolution at their coding sequence level (Patlar *et al.*, 2021), was subtracted from the transcriptomic data, and only genes with 1:1 orthology, based on Flybase orthology assignments (Thurmond *et al.*, 2019) and with a TPM greater than 1 in at least one of the two species were considered for further analyses.

After log2 transforming the expression values (Khodursky *et al.*, 2020), ANOVA was used to calculate the mean sum of squares between species (MS_{bw}) and the mean sum of squares between strains within species (MS_w). The MS_{bw} was used as a term of quantification for interspecific divergence, while the MS_w was used as a term for intraspecific variability in gene expression (Nuzhdin et al 2004; Khodursky *et al.*, 2020). Different amounts of divergence and polymorphism in expression are expected under different modes of evolution (**Table 1**) (Nuzhdin *et al.*, 2004). Genes under positive and negative selection will show a high MS_{bw} and a low MS_w, and a low MS_{bw} and low MS_w, respectively. Patterns of balancing selection are detected by a high variation within species and a low divergence between species, while genes under relaxed selection are expected to show a high variation within species and a high divergence between species (**Table 1**) (Nuzhdin *et al.*, 2004)

Mode of evolution	Polymophism (MS_w)	Divergence (MS_{bw})
Positive Selection	Low	High
Negative Selection	Low	Low
Relaxed Selection	High	High
Balancing Selection	High	Low

Table 1. **Mode of evolution of Sfp gene expression based on MS_w and MS_{bw} estimates.**

To determine whether the MS_{bw} and the MS_w was high or low for each gene in our gene set, the MS_{bw} and the MS_w estimates of each gene was compared to the estimates obtained from random gene samples of equal size to the sample of interest drawn from the rest of the transcriptome. After repeating the random sampling 10,000 times with replacement, medians were computed for MS_w and MS_{bw} (Khodursky *et al.,* 2020). A high stringency and a low stringency threshold were used as criteria to assign SFP genes into different modes of evolution. For the low stringency analysis, an estimate was considered high or low depending on whether it positioned itself above or below the genome sample midpoint, while with the high stringency analysis, estimates were considered high or low only if they were found above or below the 5% tail of the distribution (**Figure 1**).

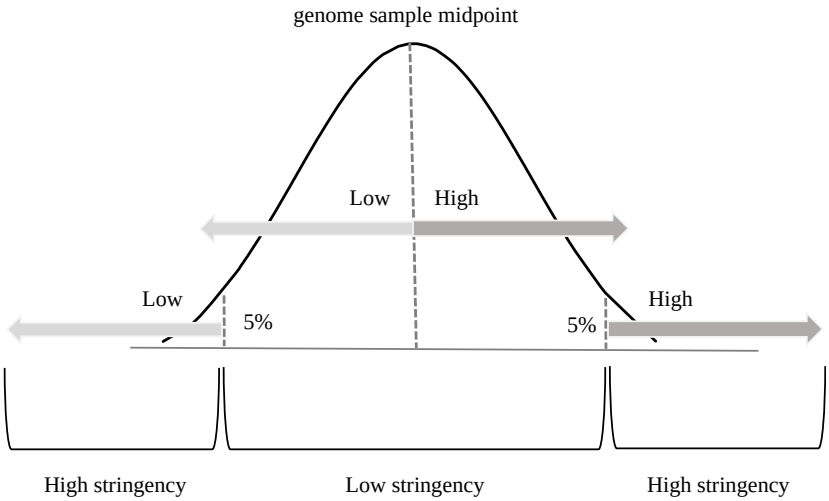

Figure 1. **Low and high stringency criteria used for SFP genes classification into different modes of evolution.**

3.4 Fly maintenance and stocks

All flies were maintained in either 50 ml bottles or vials containing 6-8 ml of a cornmeal-yeast-agar-molasses (CYAM) food media and kept on a 12-hour light-dark cycle at 22 ± 1°C. Flies were anesthetized with CO_2 for collections, but CO_2 was avoided 24 hours prior to experiments, as it was shown to increase copulation latency (Barron, 2000). To ensure virginity, females were collected every 5-7 hours. Both *D. melanogaster* and *D. simulans* flies were left to age 4-6 days prior the mating experiments.

Wild-type flies were from an isofemale line of flies collected in Winnipeg, Manitoba (Canada). Gene-specific UAS-RNAi lines were purchased from the Bloomington *Drosophila* Stock Centre (BDSC-77372, BDSC-56994, BDSC-41618, BDSC- 56016). To generate

knockdowns, a tubulin-Gal4/TM3, Sb driver stock was kindly provided by Dr. G. Findlay. As *D. melanogaster* females hardly remate to heterospecific males, we used a *D. simulans* pickpocket (ppk23) mutant stock (Δppk) (kindly provided by V. Rutta). These mutants strongly pursue *D. melanogaster* females (Seeholzer *et al.*, 2018). Moreover, a *D. simulans* curly wing stock, purchased from the National Drosophila Stock Centre (14021-0251.079), was used to introduce a dominant curly wing mutation into the *D. simulans* ppk mutant stock, by a series of backcrosses (done by Dr. A. Civetta). The resulting flies showed a curly-wing phenotype (Δppk; Cy), which facilitated the recognition of progeny sired by the *D. simulans* males.

3.5 Generation of KD males

To obtain gene-specific knockdown males, UAS-RNAi gene-target males were crossed to the tubulin-Gal4 driver females. The tubulin promoter drives ubiquitous expression of Gal4. However, the selected genes are highly tissue specific (accessory gland specific), allowing for the activation of the Gal4 system only in those tissues where the genes are highly expressed. The tubulin-Gal4 driver is on a balancer chromosome, TM3, that has a dominant marker for the expression of the stubble bristles phenotype (Sb). Experimental knock-down males were identified in the F1 by the absence of this specific phenotype. F1 males that expressed the Sb phenotype do not undergo RNAi knockdown and were used as sibling controls in the mating experiments with a common and more similar genetic background relative to the knockdown males (**Figure 2**).

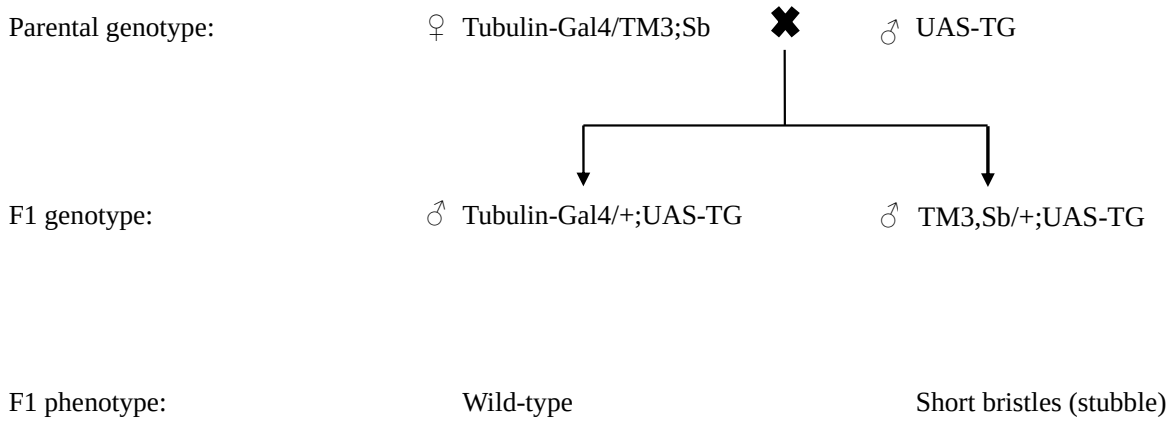

Figure 2. Crossing scheme for the generation of RNAi flies using the Tubulin-Gal4 driver. F1 progeny that lacks the stubble (Sb) phenotype has the gene-target expression knocked down via RNAi.

3.6 Confirmation of RNAi knockdowns

3.6.1 RNA extraction and cDNA synthesis

After collecting the progeny obtained from the tubulin-Gal4 – UAS-RNAi crosses, knockdown males and stubble control males were left to age 3-6 days before dissections. RNA extraction and cDNA synthesis were performed as described in the previous sections. For each gene-specific knockdown and stubble control, three biological replicates were obtained, with each sample containing RNA from 10 males. The RNA concentration and purity were then checked under the Nanodrop (ThermoFisher Scientific). Complementary DNA (cDNA) was synthetized using the iScript Select cDNA synthesis kit (Bio-Rad) from the extracted RNA.

3.6.2 Primer efficiencies and knockdown gene expression

All primers were designed using the PrimerQuest Tool available in the IDT webpage (https://www.idtdna.com/pages/tools/primerquest?returnurl=%2FPrimerquest%2FHome%2FIndex). All the primer sequences were selected to amplify products shorter than a 120 bp. To ensure the specificity of the primers, three 3-fold series dilutions were prepared from each cDNA biological replicate for each gene. A quantitative PCR (qPCR) was then performed to check the primer efficiencies. For each primer set, four reactions were prepared, one for the undiluted primer pair solution, one for the 1/3 dilution, one for the 1/9 dilution and one for the 1/27 dilution. The reactions were performed using the PowerTrack SYBR Green Master Mix (ThermoFisher) and reaction volumes were set at 20 µl, containing 10 µl of PowerTrack SYBR Green Master Mix, 1 µl of each primer pair, 6.5 µl of nuclease-free water and 2 µl of cDNA. Thermal cycling conditions were set for the first cycle at 95 °C for 2 min, followed by 40 cycles of denaturation at 95 °C for 5 s and annealing at 60 °C for 30 s. A linear equation was fit to the average quantification cycle (Cq) in relation to the dilution factor and the following equations were used to obtain primer efficiencies:

$$E = 3^{1/slope}$$

$$\% \text{ Efficiency} = (E-1) \times 100$$

Two reference genes were used to normalize gene expression, *Rps18* and *eEF1α1*. *Rps18* was previously shown to have the highest consistent expression and the least variability in expression among different genes (Grewal, 2021). The expression level of the target genes in each sample was determined by calculating ΔCq (Cycle quantification) as the Cq of the reference gene minus

the Cq of the target gene. Fold differences in relative expression were calculated using the $2^{-\Delta\Delta Ct}$ equation, where $\Delta\Delta Cq = \Delta Cq$ (control) - ΔCq (knockdown).

3.7 Sperm competition assays

Sperm competition assays at the interspecific level were performed by mating *D. melanogaster* females to a *D. melanogaster* experimental male (gene-specific knock down and control strains) and a reference *D. simulans* male. The order of mating depended on which competitive ability was being tested (sperm defense - P1 or sperm offense - P2).

For CSP assays that tested the sperm defense ability (P1) of the experimental males, the following protocol was used:

(i) Day 0: ~ 30 *D. melanogaster* virgin females and ~ 30 *D. melanogaster* experimental males (gene-specific knock down strain) were anesthetized using CO_2 and placed in a single bottle containing fly media to allow for mass mating. For each gene, a separate bottle with the same number of females and males was set up but by using siblings of the knockdown males with the Sb phenotype (non-knockdown controls). Flies were allowed to mate for ~ 24 hours.

(ii) Day 1: Flies were lightly anesthetized with CO_2 to separate males and females. Males were discarded and females were moved individually to single vials (vial 1) using an aspirator (no anesthesia).

(iii) Day 4: *D. melanogaster* females were moved individually to single vials (vial 2) using an aspirator and single *D. simulans* (Δppk; Cy) males were added to each vial to allow

for the second mating to occur. Flies were allowed to mate for 3 days before discarding the males.

(iv) Day 7: *D. simulans* males were discarded and *D. melanogaster* females were moved individually to single vials (vial 3) using an aspirator.

For CSP assays that tested the sperm offense ability (P2) of the experimental males, the following protocol was used:

(i) Day 0: ~ 30 *D. melanogaster* virgin females and ~ 30 *D. simulans* (Δppk; Cy) males were anesthetized using CO_2 and placed in a single bottle containing fly media to allow for mass mating. For each gene, a separate bottle with the same number of females and males was set up to be used for the control mating experiments. Flies were allowed to mate for ~ 24 hours.

(ii) Day 1: Flies were lightly anesthetized with CO_2 to separate males and females. Males were discarded and females were moved individually to single vials (vial 1) using an aspirator (no anesthesia).

(iii) Day 3: *D. melanogaster* females were moved individually to single vials (vial 2) using an aspirator and single *D. melanogaster* experimental males (gene-specific knockdowns) were added to each vial to allow for the second mating to occur. For each gene, the females from the control bottle were offered a single *D. melanogaster* control male (Sb) for mating.

(iv) Day 4: *D. melanogaster* experimental males (gene-specific knockdowns and controls) were discarded.

(v) Day 7: *D. melanogaster* females were moved individually to single vials (vial 3) using an aspirator.

Progeny was counted 21 days after the set-up of each vial (v1 – v2 – v3). Absence of progeny from vial 1 was used as an indication of no first mating and the vials were discarded. Absence of second male progeny from vials 2 and 3 was used as indicator of no second mating and the vials were used only for estimates of proportion of rematings. The proportion of straight-wing and curly-wing progeny was counted to estimate either P1 or P2.

3.8 Statistical analyses

All data was analyzed using RStudio (v. 4.1.1; R Core Team 2021). For every single gene, we fit an ANOVA model with expression~species + strain. Strain was taken to be a random effect. The knockdown effect on gene expression and sperm competitiveness (P1 and P2) were assessed using a one-tail Welch's t-test (H0: $\mu knockdown < \mu control$). Refractoriness and remating were tested using 2×2 Fisher Exact tests with KD or control as rows and mated or non-mated as columns. P-values were adjusted when needed to control for false positives, given multiple tests, using the Benjamini-Hochberg procedure (threshold $\alpha = 0.05$).

4.0 Results

4.1 In silico search for *D. melanogaster* SFP genes without *D. simulans* orthologs

Out of a list of 292 SFPs transferred to females during mating (Wigby *et al.*, 2020), only 20 did not have annotated orthologs in *D. simulans* (**Table 2**).

Gene ID	Gene symbol	Gene name
FBgn0004414	msopa	male-specific opa containing gene
FBgn0004426	LysC	Lysozyme C
FBgn0010357	betaTry	βTrypsin
FBgn0011668	Mst57Da	Male-specific RNA 57Da
FBgn0011669	Mst57Db	Male-specific RNA 57Db
FBgn0011670	Mst57Dc	Male-specific transcript 57Dc
FBgn0011694	EbpII	Ejaculatory bulb protein II
FBgn0023197	Jon74E	Jonah 74E
FBgn0031276	CG12506	-
FBgn0034010	CG8157	-
FBgn0051704	CG31704	-
FBgn0083936	Acp54A1	Accessory gland protein 54A1
FBgn0259963	Sfp33A2	Seminal fluid protein 33A2
FBgn0259973	Sfp79B	Seminal fluid protein 79
FBgn0261057	Sfp36F	Seminal fluid protein 36F
FBgn0261853	CG42782	-

FBgn0262623	CG43147	-
FBgn0263237	CG43319	-
FBgn0263597	Acp98AB	Accessory gland protein 98AB
FBgn0265349	Sfp33A4	Seminal fluid protein 33A4

Table 2. **List of genes that show no annotated orthologs in *D. simulans* on FlyBase**.

4.2 Gel assays for confirmation of missing orthologs in *Drosophila simulans*

To confirm the actual absence of SFP gene orthologs in *D. simulans*, RNA was extracted from the reproductive tract of both *D. melanogaster* and *D. simulans* males, and a PCR reaction followed by gel electrophoresis was used to visualize the presence of transcripts. Before running the PCR with cDNA to visualize any substantial difference in gene expression between the two species, the amplification ability of the primers was tested by running a PCR on extracted DNA from both *D. melanogaster* and *D. simulans*. Out of the 20 candidate genes, one set of primers (*CG8157*) did not show any amplification in the two species, while another one (*Acp98AB*) showed faint bands in both *D. melanogaster* and *D. simulans*. Three of them (*Sfp36F*, *betaTry*, *Mst57Da*) showed faint bands in *D. simulans*. While two of them (*Sfp33A2*, *Sfp33A4*) showed no bands in *D. simulans* (**Figure S1**). For the gene for which the primer set did not show any amplification in the two species (*CG8157*), and for two of the genes which showed no bands in *D. simulans* (*Sfp33A2*, *Sfp33A4*), the PCR with cDNA was not performed.

Out of the remaining 17 genes, only 5 showed a significant difference/lack in expression between the two species (**Figure 3**). However, Mst57Da's lower expression in *D. simulans* is likely a consequence of the low efficiency of the primers to amplify the target (see **Figure S1**). The others showed either no expression in any of the two species, a not substantial difference in expression, or a lower expression in *D. melanogaster* (**Figure S2**). Genes that showed a

lower expression in *D. melanogaster* were automatically out of our interest, as one of the methods to test the hypothesis of this thesis was to perform gene perturbation on *D. melanogaster* genes to visualize any phenotypic effect caused by the induced lower expression of the genes. Because of the small number of genes that were proved to be missing in the reproductive tract of *D. simulans*, or that have noticeable lower expression in *D. simulans*, I decided to follow a different approach for the search of genes with major differences in gene expression between the two related species.

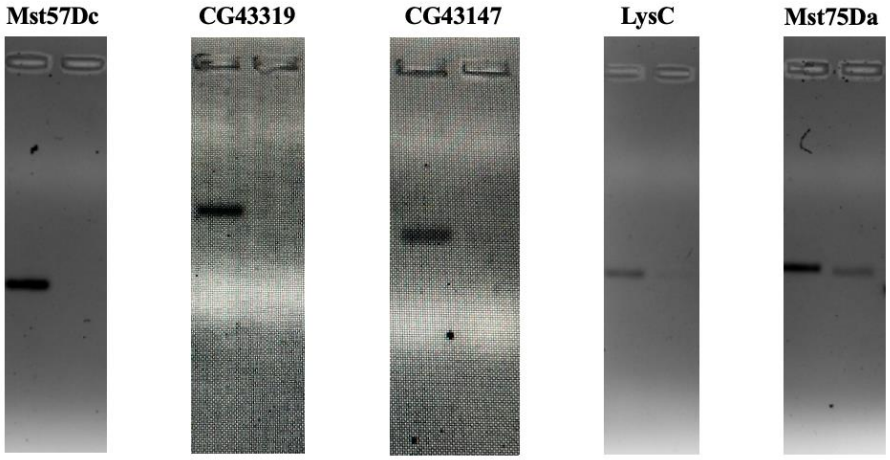

Figure 3. **Electrophoresis on agarose gel with *D. melanogaster* and *D. simulans* cDNA samples to visualize differences in gene expression.** For each gene, the *D. melanogaster* bands are the ones on the left, while the *D. simulans* bands are the ones on the right. For each of these 5 genes, the *D. simulans* band is missing, or appears significantly fainter than the *D. melanogaster* one, confirming the lower gene expression in *D. simulans*; except for *Mst57Da* due to the poor efficiency of the primers (**Figure S1**).

4.3 Transcriptomic analysis: evolutionary patterns of accessory gland genes and selection of SFPs for interspecific mating experiments

After subtracting, from the transcriptomic data, the 317 SFP genes that had previously been tested for evolutionary patterns at their coding sequence level (Patlar *et al.*, 2021), and after removing genes that lack 1:1 orthology and that did not show a TPM > 1 in at least one of the two species, 255 genes were left to analyze. Estimated of interest (MS_{bw}, MS_w and MS_{bw}/MS_w),

obtained from the genome samples, and the SFP genes were compared. P-values were obtained by comparing the median value of the SFP estimates to the 10,000 random genome samples median value, each of equal size to the SFP sample (Khodursky *et al.*, 2020). The results showed that, overall, the divergence between species (MS_{bw}) in SFP genes is higher compared to the rest of the genome (P = 0.012). However, both polymorphism (P= 0.203) and the divergence to polymorphism ratio (P=0.101) ratios of SFP genes compared to the rest of the genome showed non-significant differences (**Figure 4**). When only considering genes that encode for protein products that are known to be transferred to females during mating (Wigby *et al*., 2020), consistent results were obtained (MSbw: P=0.010; MSw: P= 0.350; and MSbw/ MSw: P= 0.101).

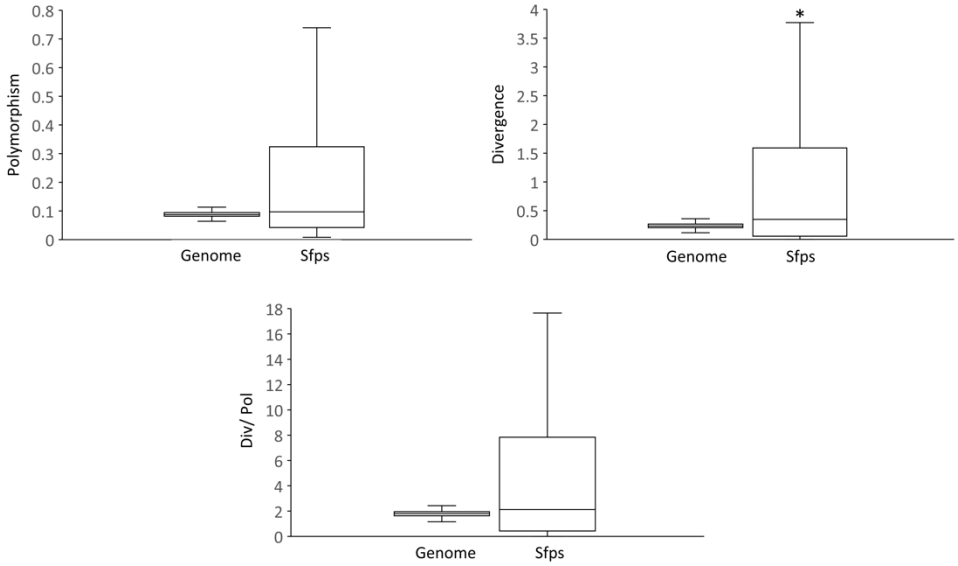

Figure 4. **Divergence (interspecific), polymorphism (intraspecific) and divergence to polymorphism ratio (interspecific to intraspecific variability) ratios in gene expression in the male accessory gland transcriptome.** SFP genes are compared to the estimates of the rest of the genome. (* P-value < 0.05).

In order to obtain a SFP gene classification based on their mode of evolution, two different criteria were used to establish whether each SFP gene was following a pattern of high

or low variation within species and high or low divergence between species. The first criterion assigned the gene estimate a high or low value based on whether the estimate was lower or higher than the midpoint estimate obtained from the genome random sampling. The other criterion used a more stringent method of classification. Genes were classified based on the position of the gene estimate compared to the 5% tail of the genome distribution (**Figure 5**).

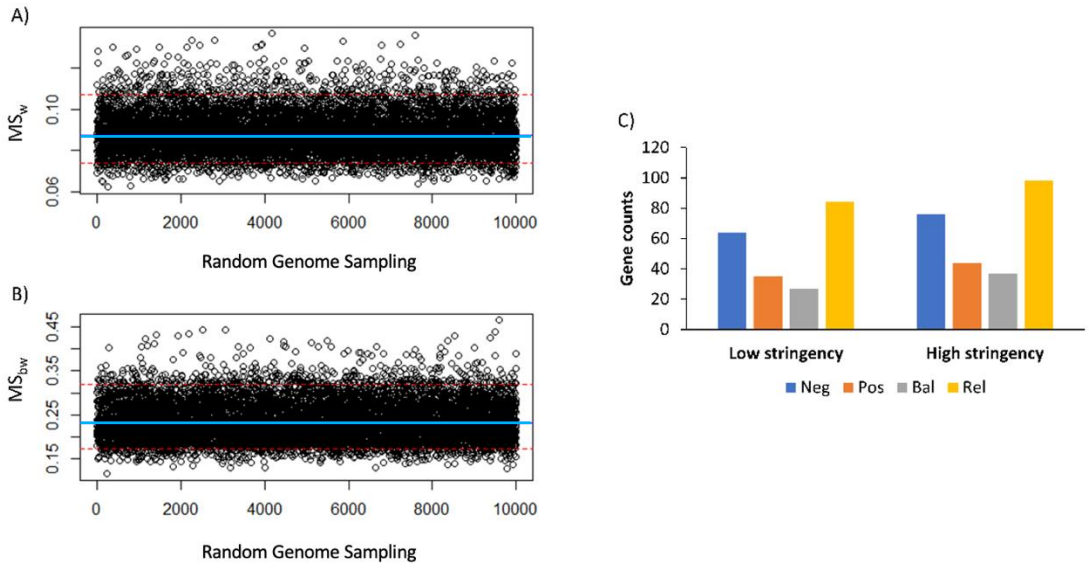

Figure 5. **Mode of evolution of SFP genes.** Figure A is showing the genome mean square of estimates of polymorphism within species (MS_w), while figure B is showing the mean square of estimates of divergence between species (MS_{bw}). The low stringency classification that identified an estimate value as low or high was based on the median as threshold, indicated by the blue solid line, while the red dashed lines indicate the upper and lower 5% tails thresholds used in the high stringency classification. Figure C shows the counts of SFP genes under the four different modes of evolution. (Neg= negative selection - low variation, low divergence; Pos= Positive selection - low variation, high divergence; Bal= Balancing selection - high variation, low divergence; Rel= Relaxed selection - high variation, high divergence).

The results of this classification identified a larger proportion of SFP genes as highly diverged between species (57% vs. 43%; $\chi 2$= 6.94; P= 0.0084), while showing no difference in the proportion of SFP genes displaying low variation vs high variation in expression (47%

vs 53%; $\chi 2$= 1.15; P= 0.2831). When the less stringent cut off was used, the results were consistent ($\chi 2$= 6.15; P= 0.0132; $\chi 2$= 1.54; P= 0.215; respectively).

Patlar *et al.* (2021), found that 57% of their list of 317 SFP genes from the Raleigh population, analyzed for patterns of evolution at their coding sequence levels, were under relaxed selection, with only ~7% of genes being under positive selection, and 36% under selective constraints. My expression data results, irrespective of the used criterion, showed that 38.4 - 40.0% of SFP genes are relaxed. However, compared to coding sequence analysis results, there was a significant increase in the proportion of positive selected genes (17%) (6.6% to 17%; $\chi 2$= 10.7, P = 0.0011), with no difference in the proportion of negative selected genes (29.8 - 30.5%) (36% to 30%; $\chi 2$= 1.3, P = 0.2525). These results show that, similar to coding sequence data (Patlar *et al.*, 2021), the high divergence in expression between species in SFP genes is attributable to relaxed selection.

Genes under positive selection that affect ISC are prime candidates for cases of post copulatory sexual selection driving speciation. In the classification obtained from the transcriptomic analysis, four genes that resulted to be under positive selection (**Table 3**) had been previously tested and confirmed to have a role in ISC (Patlar and Civetta, 2022, Civetta and Ranz, 2019).

Gene name	Mode of Evolution	ISC	Role
Acp29AB	Positive directional	YES	**P1***, **sperm storage** (Findley *et al.*, 2008; Mueller *et al.*, 2005; Wong *et al.*, 2008; Patlar and Civetta, 2022)
Acp62F	Positive directional	YES	**P1** (Findley *et al.*, 2008; Mueller *et al.*, 2005; Mueller *et al.*, 2008 Wong *et al.*, 2008; Civetta and Ranz, 2019)
lectin-46Ca	Positive directional	YES	**P1**, **female receptivity** (Ram and Wolfner, 2007; Avila and Wolfner, 2009; Civetta and Ranz, 2019)
CG9168	Positive directional	YES	**P1** (Patlar and Civetta, 2019)

Table 3. **Candidate genes for interspecific mating experiments known to have roles in intraspecific sperm competition.**
*** "Role in P1" indicates that the gene's protein product was shown to have a role in the sperm defense mechanism, allowing stored sperm to resist displacement from rival sperm in sperm competition settings.**

Moreover, these four genes form a network of interaction, with *Acp29AB* being the central hub (**Figure 6**).

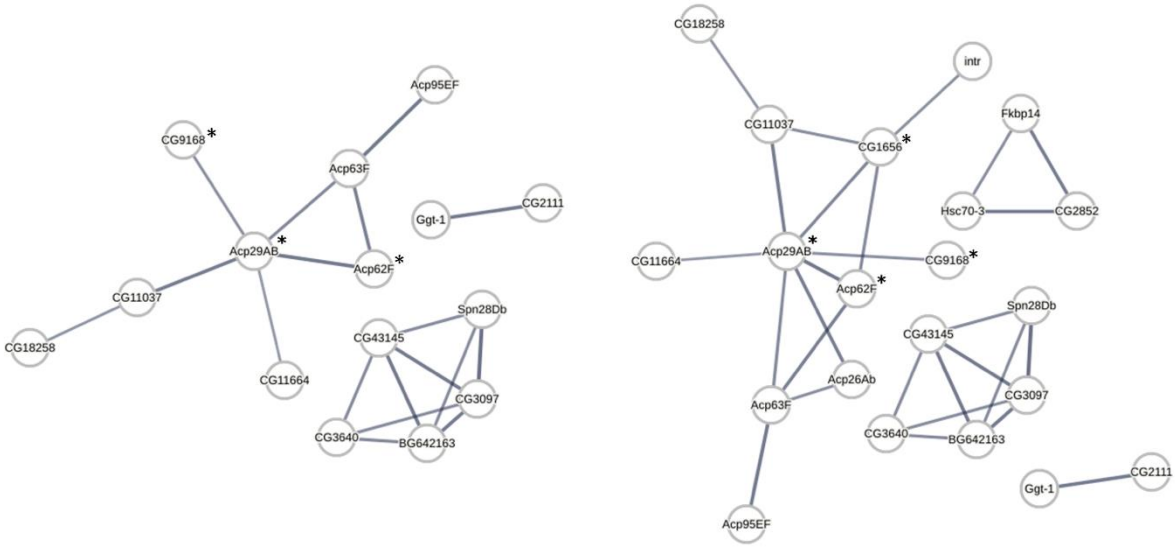

Figure 6. **Protein-protein interaction (PPI) network for genes under positive selection known to affect ISC**. The PPI was generated using STRING (https://string-db.org/). On the left, the figure shows the network with high confidence interaction scores, while on the right, medium confidence interaction scores were used. The four genes known to affect ISC are highlighted using an Asterix, with *CG1656* being *lectin-46Ca*.

It is known that interactions between seminal fluid proteins are required to generate long term post-mating responses (Chapman *et al.*, 2003; Liu and Kubli, 2003). An example is the "sex-peptide network", which shapes female's remating, egg production and mating behaviours. The induction of the post-mating responses relies on the interaction of *sex-peptide* with other conserved seminal fluid proteins, and it was shown that lack of some of these proteins would result in the absence of parts of the post-mating responses (Ravi Ram and Wolfner, 2007). For this reason, I hypothesize that the knockdown of the four positively selected genes will decrease the male's ability to outcompete a heterospecific rival in mating experiments. Therefore, I focused on these four genes to perform mating experiments to test the male paternity success in sperm competition at the interspecific level, between *D. melanogaster* KD males and *D. simulans* males. This allows to test whether these genes have a role in CSP, which would provide evidence of a common genetic basis between forms of post-copulatory sexual selection and reproductive isolation.

4.4 Genomic and functional features as predictors of SFP gene expression evolution

In order to get a better understanding of SFP genes' mode of evolution, an analysis of association between assigned selection categories and genomic or functional features (i.e., sex vs autosomal location, tissue of expression, post mating effects, immunity, proteases – Data from Patlar *et al*., (2021)) was performed. I performed the association analysis using the more stringent criteria for classification but results under less stringent analysis are shown when in disagreement. The results showed a non-random association between chromosome location

and SFP gene mode of evolution, with a marginal underrepresentation of genes under relaxed selection on the X-chromosome (P= 0.067) (**Table 4**). The results become significant under a less stringent criterion (Padj=0.047; Odds ratio= 0.230). Moreover, the association analysis showed an enrichment of positively and negatively selected genes among genes with specificity of expression in the male's reproductive tract, with an underrepresentation of relaxed genes (**Table 4**). Among genes with reproductive tissue-specificity, AG-SFP genes had an excess of genes under positive and negative selection (**Table 5**). However, when a less stringent criterion was used, negative selected genes were not enriched among genes with specificity of expression in reproductive tissues (Padj=0.140; Odds ratio= 1.675), or they showed only a marginal effect for AG-SFP (Padj=0.058; Odds ratio= 1.888) (**Table 5**). Instead, genes with other tissue-specificity of expression or non-tissue specific were under relaxed selection (**Table 5**).

Among genes that had previously been shown to affect intraspecific sperm competition *via* gene knockdown, the most represented mode of evolution was found to be negative selection (FET: P= 0.0573, **Table 6**). Testing of gene-knockdown on both intraspecific and interspecific sperm competition has only been done for three genes (*SP*, *CG9997*, *Acp36DE*) (Castillo and Moyle, 2014). Interestingly, the two genes that showed to affect both ISC and CSP appear to be under negative selection, while *SP*, which was only found to affect ISC, shows a pattern of relaxed selection (**Table 6**). My results identify four genes (**Table 6**) whose changes in expression appear to be driven by positive selection.

		Selection Regime							
		Negative		Positive		Balancing		Relaxed	
Feature	**P-value ***	**Odds Ratio**	**P-adj †**	**Odds Ratio**	**P-adj †**	**Odds Ratio**	**P-adj †**	**Odds Ratio**	**P-adj †**
Transferred vs. non-transferred	0.232	0.802	0.849	4.866	0.376	0.482	0.376	1.041	1.000
X vs. Autosomes	**0.030**	1.255	0.728	3.093	0.102	1.219	0.728	0.239	0.067↓
Reproductive vs. non-reproductive	**6.6E-4**	2.151↑	**0.035**	3.442↑	**0.020**	0.676	0.383	0.362↓	**0.004**
Post-mating vs. unknown	0.114	1.340	0.509	0.656	0.509	0.239	0.193	1.520	0.482
Immunity vs. unknown	0.887	0.849	1.000	0.487	1.000	0.667	1.000	1.856	1.000
Proteases vs. non-proteases	0.407	0.786	0.822	1.916	0.645	0.283	0.645	1.174	0.822

Table 4. **Patterns of non-random association for six genomic or functional features and different gene categories based on their mode of molecular evolution.**
* For each feature, genes are split into two categories. The P-value is bolded if significant. The arrow identifies the excess or deficit for the first category listed. For the alternative category the pattern is the opposite (e.g., reproductive SFPs are underrepresented, while nonreproductive are enriched, in the relaxed selection group). † Post hoc 2x2 FETs to test for significant excess (odds ratio>1) or deficit (odds ratio<1) between any selective regime and the others. P-values are FDR corrected.

		Selection Regime							
		Negative		Positive		Balancing		Relaxed	
Comparison	***P*-value ***	**Odds Ratio**	***P*-adj †**	**Odds Ratio**	***P*-adj †**	**Odds Ratio**	***P*-adj †**	**Odds Ratio**	***P*-adj †**
High stringency AG-specific *vs*. others	**2.4E-5**	2.519↑	**0.010**	4.035↑	**0.007**	0.634	0.390	0.293↓	**2.1E-4**
Low stringency AG-specific *vs*. others	**8.7E-5**	1.888↑	0.058	3.372↑	**0.006**	0.995	1.000	0.322↓	**1.7E-4**

Table 5. **Pattern of non-random association of accessory gland tissue-specific genes with different modes of evolution**. * For each comparison, genes are split into two categories. A 2x4 Fischer Exact Test (FET) is used to evaluate the differential association with the four modes of evolution. † Post hoc 2x2 FETs is used to test for significant excess (odds ratio>1) or deficit (odds ratio<1) between any selective regime and the others. P-values are FDR corrected. The P-value is bolded if significant. The arrow identifies the excess or deficit for the first category listed in the comparison. For the alternative category, the pattern is the opposite.

Gene	Tissue	MS_w*	MS_{bw}*	Mode evolution† (stringent)	Mode evolution (relaxed)	ISC^	CSP^
CG11598	AG	0.0427↓	0.0003↓	negative	negative	yes	NA
Acp76A	AG	0.0629↓	0.0345↓	negative	negative	yes	NA
Qsox4	AG	0.0231↓	0.0115↓	negative	negative	yes	NA
aqrs	AG	0.0359↓	0.0018↓	negative	negative	yes	NA
CG9997	AG	0.0259↓	0.1007↓	negative	negative	yes	yes
CG17575	AG	0.0528↓	3.0E-05↓	negative	negative	yes	NA
Acp26Aa	AG	0.0420↓	0.0085↓	negative	negative	yes	NA
Acp36DE	AG	0.0754↓	0.0419↓	NA	negative	yes	yes
lectin-46Cb	AG	0.0433↓	0.2017↓	NA	negative	yes	NA
CG9168	AG	0.0157↓	0.4902↑	positive	positive	yes	NA
Acp29AB	AG	0.0492↓	1.0734↑	positive	positive	yes	NA
Acp62F	AG	0.0207↓	0.8912↑	positive	positive	yes	NA
lectin-46Ca	AG	0.0215↓	0.3100↑	NA	positive	yes	NA
Acp53Ea	AG	0.0888↑	0.0010↓	NA	balancing	yes	NA
Acp33A	AG	0.1445↑	2.5503↑	relax	relax	yes	NA
Semp1	AG	3.6580↑	2.4887↑	relax	relax	yes	NA
CG17242	AG	0.1624↑	5.2904↑	relax	relax	yes	NA
Est-6	ED	0.1179↑	0.3460↑	relax	relax	yes	NA
SP	AG	0.1058↑	0.5718↑	NA	relax	yes	no

Table 6. **Variation and divergence in expression for SFP genes involved in sperm competition**. * Arrows identifies a value lower (down) or higher (up) than the genome sample midpoint. † Under a stringent classification, only genes with estimates beyond the lower or upper 5% tail of the genome distribution are classified. ^ Data from Civetta and Ranz 2019; and Patlar and Civetta 2022. NA= not available.

4.5 Gene knock-down confirmation

4.5.1 Primer efficiencies

Primers for the four genes of interest and the two reference genes used to normalize the data (*Rps18*, *eEF1α1*) were designed and tested for efficiency before performing a qPCR to confirm the downregulation of gene expression in the knockdown flies. After plotting a linear regression curve and calculating the slope of the trend line (**Figure 7**), all primers showed adequate efficiency to perform qPCR analysis (**Table 7**). Ideally, all the primers should have an efficiency that ranges from 90% to a 100%. However, an efficiency that slightly exceeds 100% (~ 110%) can also be accepted.

Gene	Primer sequence	Amplicon size	Primer efficiency (%)
Acp29AB	R: TGGAGTTTAAGGCCCAGATG F: GATGTTGGATGCATGGTGTC	98	94
Acp62F	R: TGATAACATATCCCGGCTTAC F: GTCCTGTAGCATGTCCTGAAA	99	93
CG9168	R: GGACTGAAATCCGCCATAGAA F: ACCTGAAGATCAACCAGAATCG	103	112
lectin-46Ca	R: TGATGCCCACATAGAAGCAC F: CGGAAAGAAGCAGAGCAAA	95	112
RpS18	R: GATCGACATCGGCCTTCTTC F: CAGCACATCCTGCGTATCAT	127	103
eEF1α1	R: GCTGCTGTTGCTGGTATTATTG F: TAGCTGGTTTGCTTCCACTC	120	91

Table 7. **Primer sequences and primer efficiency for the four genes of interest** (*Acp29AB*, *Acp62F*, *CG9168*, *lectin-46Ca*) **and the reference genes** (*RpS18* and *eEF1α1*) **used to test the effectiveness of the GAL4 system in the downregulation of gene expression in the accessory glands of *D. melanogaster* males.**

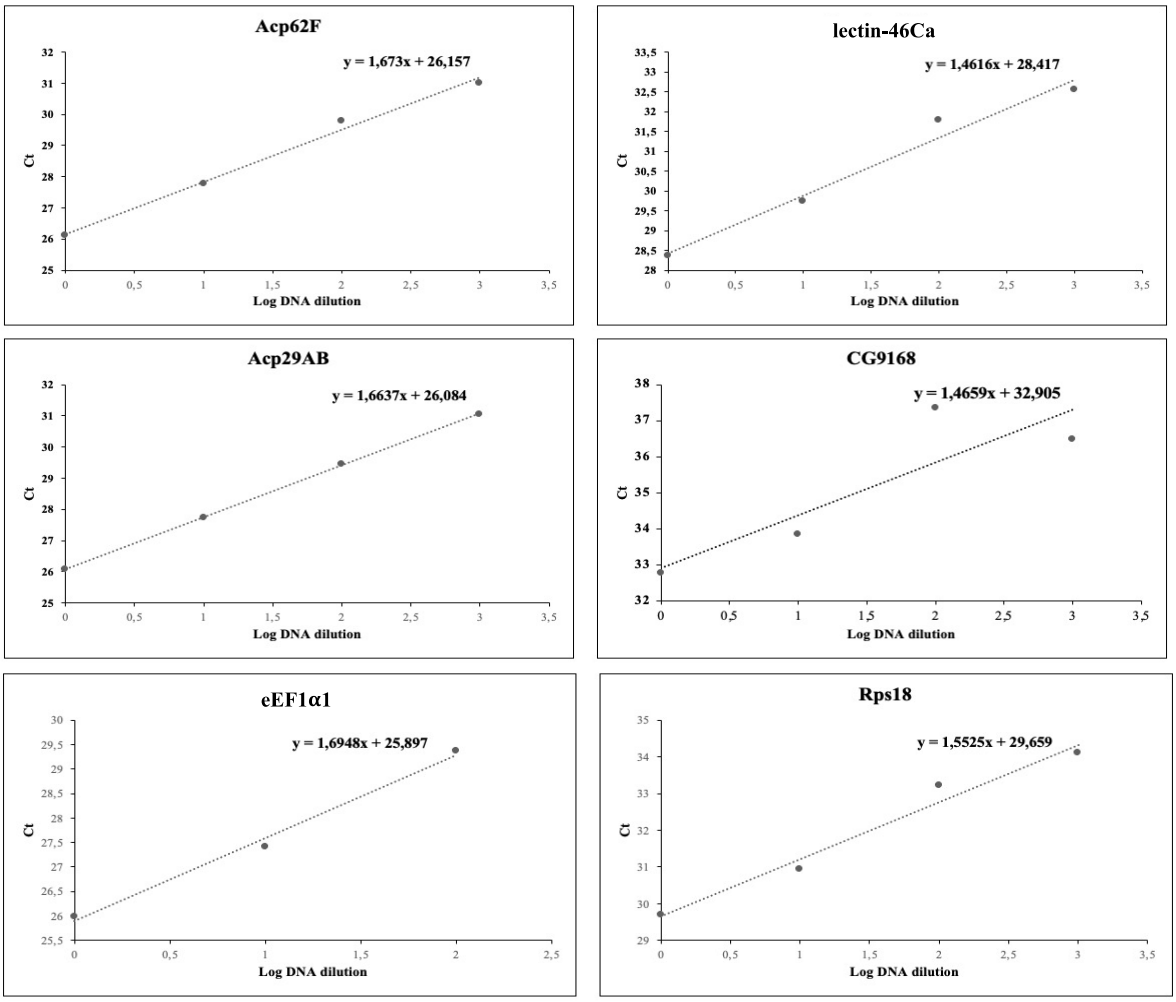

Figure 7. **Slope of linear regression fit for primer efficiency.**

4.5.2 Gene knockdown verification

The four selected genes (*Acp29AB, CG9168, Acp62F, lectin-46Ca*) were knocked down by using the tubulin-Gal4 system. As previously mentioned, this system also generates a sibling control of the knockdowns, that is visible due to the stubble bristle phenotype (Sb). Sb control flies lack the Gal4 driver and are not expected to experience a knock down of the targeted gene expression. To verify the correct functioning of the tubulin-Gal4 system, a qPCR was performed by comparing the expression of the gene-specific knockdown flies to the expression of their specific Sb sibling controls. qPCR was normalized using both the housekeeping genes *Rps18* and *eEF1α1*. The results showed that all the KDs had a significant reduction of expression for the four genes in comparison to their sibling Sb controls (**Figure 8**).

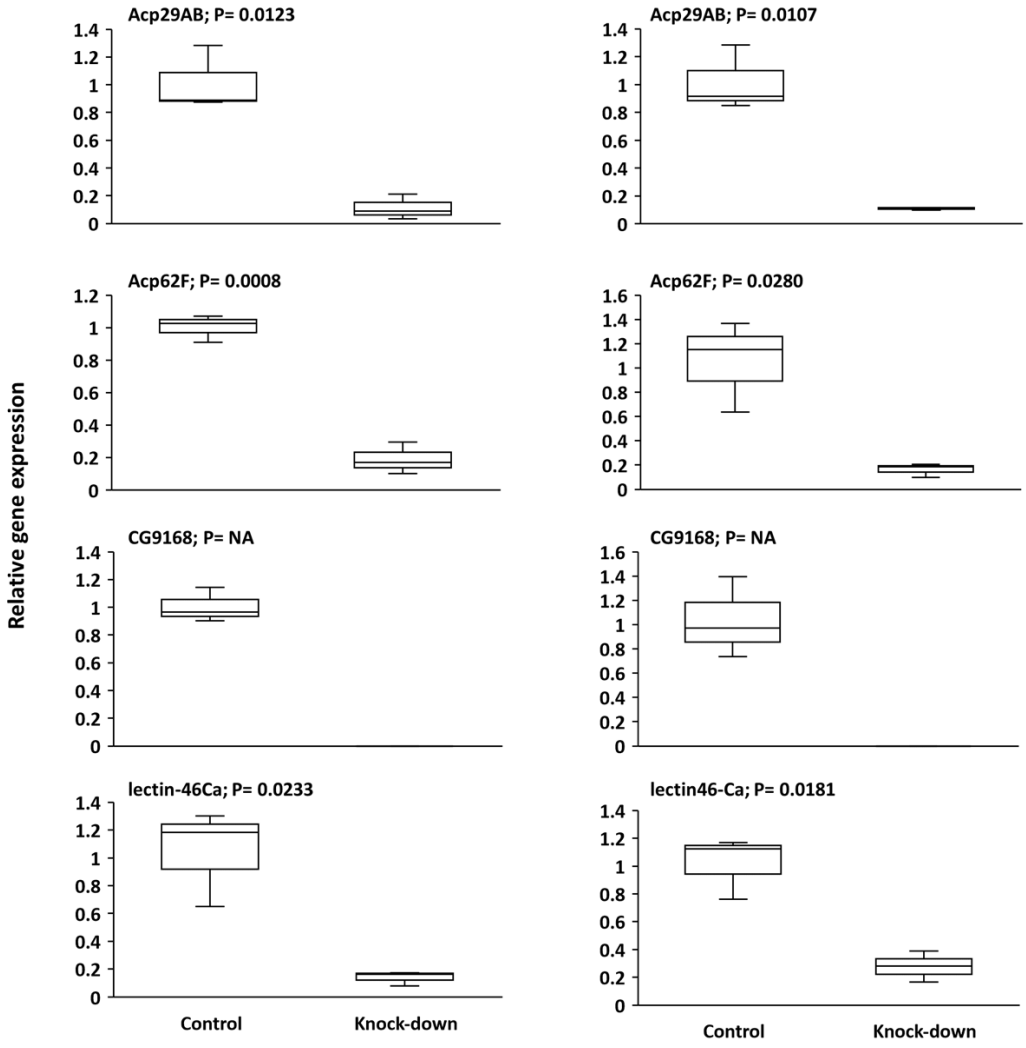

Figure 8. **RNAi knockdown efficacy**. To compare the average relative gene expression of the knockdowns compared to their sibling Sb controls, a one-tailed two samples Welch's test was performed. on the left, the relative gene's expression is normalized with *Rps18*. On the right, the relative gene expression in normalized with *eEF1α1*. For *CG9168*, the knockdown's gene expression was non- detectable, when normalized with both housekeeping genes. However, as the sibling control's gene expression resulted normal, it indicated a nearly complete down-regulation of the gene.

4.6 Sperm competition

Both the offensive and defensive ability of the knockdown males were measured by counting the amount of progeny sired by the *D. melanogaster* KD males and the heterospecific *D. simulans* curly wing males. To test the P1 ability of the knockdowns, KD males were used

for the first mating, while *D. simulans* males were used in the second mating. To test the P2 ability of the knockdowns, *D. simulans* males were used in the first mating, while the KD males were used in the second mating. Because in these heterospecific crosses, only female offspring is viable, female-only progeny was used to calculate the estimates for P1 and P2. The resulting values of P1 and P2 of the experimental conditions were compared to the P1 and P2 values that resulted from the control mating experiments.

Mating experiments that tested the defensive ability (P1) of the KD males showed that none of the gene-specific knockdowns affected the conspecific KD males ability to outcompete their heterospecific rivals (**Table 8**) (**Figure 9**). Similarly, mating experiments that tested the offensive ability (P2) of the knockdown males showed that the ability to sire progeny was not affected by the gene-specific knockdowns when the KD males were mated second (**Table 9**) (**Figure 9**). The results are consistent when both male and female progeny are included in the estimates of P1 and P2 (**Tables 8** and **9**).

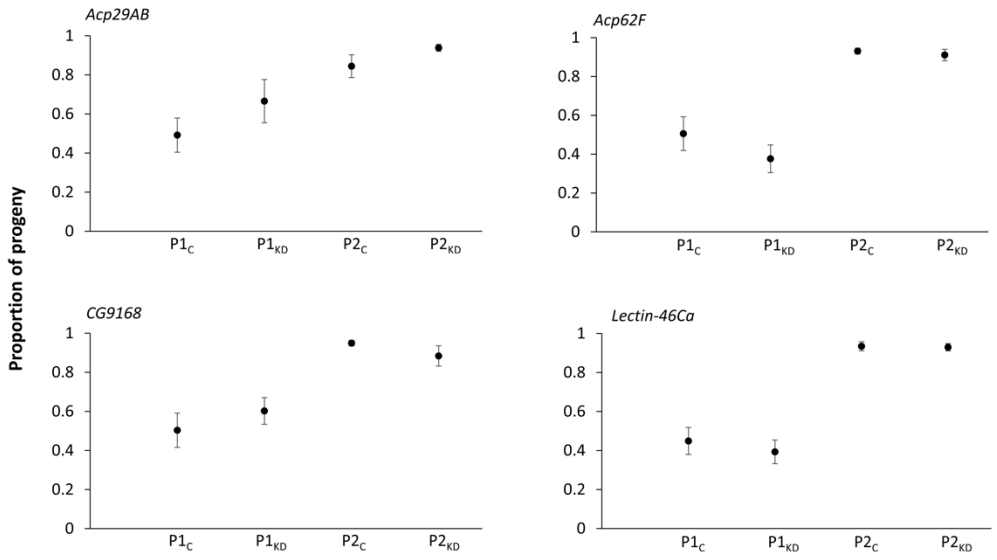

Figure 9**. Proportion of progeny sired by each gene's KD males when first (P1) and second (P2) to mate with a *D. melanogaster* female, compared to their sibling Sb control males**

(P1c and P2c). In all the tests, the competitor is a *D. simulans* male. Each point represents mean ± s.e. of the mean.

		Females only				All			
Gene	Treatment	N	P1	*t*	*P*	N	P1	*t*	*P*
Acp29AB	Control	14	0.49±0.33	1.23	0.1159	14	0.56±0.32	1.61	0.1159
	KD	11	0.67±0.37			11	0.76±0.29		
Acp62F	Control	16	0.51±0.35	1.15	0.1299	16	0.59±0.32	1.06	0.1491
	KD	17	0.38±0.29			17	0.48±0.28		
CG9168	Control	12	0.50±0.30	0.67	0.2561	12	0.61±0.26	1.15	0.1311
	KD	16	0.60±0.27			16	0.72±0.23		
lectin-46Ca	Control	26	0.45±0.35	0.60	0.2742	26	0.53±0.35	0.73	0.2355
	KD	31	0.39±0.34			31	0.46±0.34		

Table 8. **Conspecific sperm defense ability (P1) of knockdown (KD) versus Sb sibling control (no-KD) males**. One-tailed Welch's t-tests were performed to test mean differences for the proportion of offspring sired by knockdown and control males. Only female progeny from the heterospecific cross is viable so we used females-only progeny for our estimates. Results are consistent if males are also included in the analysis (All). N= Total number of males tested, P1= Mean ± Standard deviation of the sample mean.

		Females only				All			
Gene	Treatment	N	P2	*t*	*P*	N	P2	*t*	*P*
Acp29AB	Control	16	0.84±0.23	-1.54	0.0709	16	0.89±0.24	-1.31	0.1056
	KD	21	0.94±0.08			21	0.97±0.04		
Acp62F	Control	33	0.93±0.08	0.61	0.2708	33	0.59±0.32	0.70	0.2444
	KD	35	0.91±0.17			35	0.48±0.28		
CG9168	Control	19	0.95±0.06	1.22	0.1193	19	0.97±0.03	1.10	0.1433
	KD	18	0.88±0.22			18	0.91±0.22		
lectin-46Ca	Control	16	0.93±0.09	0.17	0.4338	16	0.96±0.06	0.15	0.4399
	KD	15	0.93±0.07			15	0.96±0.05		

Table 9. **Conspecific sperm offense ability (P2) of knockdown (KD) versus Sb sibling control (no-KD) males**. One-tailed Welch's t-tests were performed to test mean differences for the proportion of offspring sired by knockdown and control males. Only female progeny from the heterospecific cross is viable so we used females-only progeny for our estimates. Results are consistent if males are also included in the analysis (All). N= Total number of males tested, P1= Mean ± Standard deviation of the sample mean.

4.7 Female refractoriness and remating

Drosophila melanogaster is a polyandrous species, and females accept to mate with more than one male in a short period of time. However, females of this species have a refractory period after mating with the first male where they refuse to mate to a second male. Some seminal fluid proteins (e.g. *SP*) are known to extend the female's refractory period in order to prevent a second mating to occur and sire most of the progeny (Chapman *et al.*, 2003; Liu and Kubli, 2003). The genetic manipulation of these proteins can induce a female to remate with a second male more quickly, and increases the chances of sperm competition (Leiblich, Marsden, Gandy *et al.*, 2012; Hopkins, Sepil *et al.*, 2019). For this reason, it was assumed that some of the four genes of interest might have a role in female's refractoriness and female's remating.

The results of the heterospecific mating experiments, however, showed that when females mated first with the KD males for each of the four genes, they were not less reluctant to remate to the *D. simulans* males compared to their sibling controls, suggesting a non-involvement of these genes in the female's refractory period. Moreover, the results of the offensive tests showed that, for all the four genes, KD males did not have a decreased ability to mate to females after they had already mated to the *D. simulans* heterospecific males (**Table 10**).

		Refractoriness			Remating		
Gene	Treatment	N	Odds	*P*	N	Odds	*P*
Acp29AB	Control	14/23	1.38	0.75	16/19	1.94	0.64
	KD	13/19			21/23		
Acp62F	Control	16/27	1.06	1.00	33/37	2.10	0.67
	KD	17/28			35/37		
CG9168	Control	12/25	1.43	0.59	19/22	1.41	1.00
	KD	16/28			18/20		
lectin-46Ca	Control	26/35	2.65	0.22	16/18	0.94	1.00
	KD	31/35			15/17		

Table 10. **Females' refractoriness (P1) and remating (P2).** Fisher Exact tests were performed to compare knockdown and control males for female refractoriness to remate to an heterospecific male after mating to a conspecific male. Remating compares knockdown and control males for the proportion of females that mate to conspecifics after a first mating to an heterospecific male. (N: Number of females remated / Number of females that did not remate).

5.0 Discussion

After performing PCR experiments to confirm the lack of orthology between SFP genes between *D. melanogaster* and *D. simulans*, as predicted by the Flybase database, we found that only four out of the 20 selected genes were lacking or had reduced expression in *D. simulans*. Orthologous genes share a common genetic ancestor as they are inherited copies of a single gene after a speciation event. Online databases use sequence similarity, segment synteny (shared location within the chromosome between different species) and functional domains to determine orthologous genes (Vallender *et al.*, 2009). However, bioinformatic tools are not infallible, and the finding of false negatives and positives has often occurred (Vallender *et al*., 2009). In this thesis, false negatives (16 out of 20) seem to be the most probable explanation. There are different scenarios that could explain errors in the identification of orthologs when using bioinformatic tools. They could be a consequence of difficulties in genome annotations, errors in *de novo* gene predictions, identification of false transcripts and sequencing errors (Vallender, 2009).

Interestingly, among the 20 genes surveyed, is a cluster of three genes (*Mst57Dc*, *Mst57Da* and *Mst57Db).* These three genes produce RNA transcripts specific to the *D. melanogaster* male's accessory gland, and they are encoded within genome proximity (Simmerl et al., 1995). In *D. melanogaster*, all these gene protein products are transferred to the females during mating (Wigby et al., 2020). One of these genes (*Mst57Dc*) was confirmed to lack an ortholog with expression in the *D. simulans* reproductive tract. *Mst57Da*'s lower expression in D. simulans appears to be a consequence of the lower efficiency of the primers to amplify the gene in that species, while *Mst57Db* was a false negative, and showed normal expression in the reproductive tract of *D. simulans*. Gene products often interact with others forming networks, and they work together in order to generate a molecular response. To

understand whether this cluster of genes acts in a network or has any kind of relationship between them, I used data from a functional protein association networks database (STRING). Data available on predicted interactions showed, under both a medium and a high confidence interaction score, a co-expression level of 0.840 for only two of them (*Mst57Da* and *Mst57Dc*), while not showing any level of interaction with *Mst57Db*. In *D. melanogaster*, the *Mst57Db's* transcript is the most abundant of the three genes (Simmerl *et al*., 1995). A possible explanation for the absence of *Mst57* orthologs in *D. simulans* is that *Mst57Db*, in this species, has acquired the functions of the three genes with the two other transcript sequences being lost. Gene loss can be beneficial to adapt to different environmental conditions (Helsen *et al.*, 2020). However, another possible explanation of why *Mst57Db* is the only ortholog that has been retained in *D. simulans*, is that *Mst57Da* and *Mst57Dc* are newly evolved genes in *D. melanogaster*. *De novo* genes can arise from ancestrally non-coding genomic regions, and they were shown to commonly have male-biased specificity of expression, such as testis or accessory gland-specificity of expression (Levine *et al.*, 2006; Begun *et al.*, 2007; Palmieri *et al.*, 2014). To try to understand which of these two explanations might be correct, I looked at whether the closest species to *D. melanogaster* and *D. simulans* had orthologs of these genes. The closest species to *D. melanogaster* and D. *simulans*, by looking at the phylogenetic tree of *Drosophila*, are, respectively, *D. yakuba*, *D. mauritiana* and *D. sechellia* (**Figure 10**). After searching for orthologs (via orthoDB), *Mst57Da* does not show any annotated orthologs in neither *D. yakuba*, *D. mauritiana* nor *D. sechellia*, suggesting that this gene might be the result of a gene gain in *D. melanogaster*. However, *Mst57Dc* shows an ortholog in both *D. mauritiana* and *D. sechellia*, suggesting that this gene might have been lost in *D. simulans*.

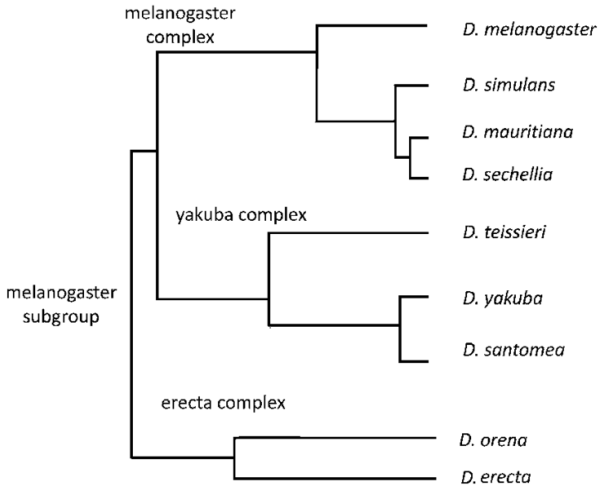

Figure 10. **Phylogenetic tree of the melanogaster subgroup of *Drosophila*** (Ambrosi *et al.*, 2013).

Future research and genetic manipulation of these genes in *D. melanogaster* might provide important insights on the evolution of SFP genes between closely related species and how gene loss/gene gain might be involved in establishing genetic and phenotypic variations between species.

Similarly to what was found in the recent study conducted on protein coding sequences (Patlar *et al.*, 2021), the results of the transcriptomic analysis showed that SFP gene expression has a higher divergence between species compared to the rest of the genome. This result agrees with many studies that have previously observed that male-biased genes show faster interspecies expression divergence (Meiklejohn *et al.*, 2003; Parisi *et al.*, 2004; Ellegren and Parsch 2007; Brawand *et al.*, 2011; Assis *et al.*, 2012; Cridland *et al.*, 2020). The high interspecific expression divergence of male-biased genes has also been observed in other species, such as between humans and chimpanzee and mouse species (Khaitovich *et al*., 2005; Voolstra *et al*., 2007). Because of their high interspecific divergence, these genes were thought to evolve under adaptive selection. However, our results show that the high expression divergence of SFP genes has been mainly driven by the relaxation of selective pressures.

Relaxed selection can be an accurate explanation of the fast expression divergence of these genes, as sex-biased genes experience a much lower selection in the sex where they are not expressed in, which is about half of the population (Pröschel *et al.*, 2006; Dapper and Wade 2020). Moreover, spatial and temporal fluctuation in the induction of gene expression were shown to favour relaxed selection (Kawecki *et al.*, 1997; Van Dyken and Wade 2010). The amount in expression of some SFPs, in fact, was observed to depend on the social environment conditions in *D. melanogaster* males, in response to the intensity of male-male competition (Fedorka *et al.*, 2011).

Phenotypic evolution is affected by mutation at both the protein coding sequence level and at the gene expression level. While agreeing with the high expression divergence result and relaxed selection being the main driver of evolution of SFP genes, the results of our transcriptomic analysis reported a significantly greater number of genes evolving under positive selection, in comparison to coding sequence results (Patlar *et al.*, 2021). Interestingly, only three of these genes (*Acp29AB*, *CG2111* and *Ggt-1*) were under positive selection for both the analysis at the gene expression level and at the coding sequence level. This limited overlap agrees on one hand, with previous evidence that suggests that mutations at the coding sequence level and at the gene expression level are responsible for the evolution of different types of genes (Wray, 2007; Haygood *et al.*, 2010; Liao *et al.*, 2010). For example, mutations at the coding sequence level have been found to mainly shape the evolution of genes involved in physiological changes, while those at the gene expression level primarily affect the evolution of genes involved in morphological traits (Wray, 2007; Haygood *et al.*, 2010; Liao *et al.*, 2010). On the other hand, however, several studies have reported a positive correlation between gene expression divergence and protein sequence evolution (Nuzhdin *et al.*, 2004; Lemos *et al.*, 2005; Khaitovich *et al.*, 2005; Sartor *et al.*, 2006; Artieri *et al.*, 2007; Hunt *et al.*, 2013; Warnefors and Kaessmann 2013; Hodgins *et al.*, 2016; Go and Civetta 2020; Zhong, Lundberg,

and Råberg 2021). If the gene expression divergence and protein sequence divergence are not correlated, this might provide positive selection with more opportunities to drive species-specific adaptations through changes at the gene expression and protein coding level.

The results of our analysis of association between the mode of evolution of SFP genes' evolution and genomic or functional features indicated that male-biased genes were enriched with genes under both positive and negative selection. This is not in agreement with the results obtained from the coding sequence analysis, where genes with male-specific tissue of expression were, instead, mostly under relaxed selection (Patlar *et al.*, 2021). The narrow expression of sex-biased genes could imply that they are less pleiotropic than non-sex-biased genes, and as they have a lower number of functional constraints, one would expect them to be evolving at a faster rate due to relaxed purifying selection (Mank *et al.*, 2008). Weakened selection on such genes could increase the segregation of slightly deleterious variation, which will cause an increase in polymorphism and a faster evolution of the genes (Van Dyken and Wade 2010; Purandare *et al.*, 2014) There have been contradictive results in terms of the effect of pleiotropy on selection efficacy (Jordan *et al.*, 2003; Hahn *et al.*, 2004; Hahn and Kern 2005; Papakostas *et al.*, 2014; Vedanayagam and Garrigan 2015; Huber *et al.*, 2017; Josephs *et al.*, 2017), and without a direct test on fitness effects, breadth of expression cannot be directly associated with true pleiotropy. Hence, it is possible that genes with tissue-specific expression might be enriched for positive and negative selection, given their essential functions in reproduction (Ravi Ram and Wolfner 2007; Civetta and Ranz 2019; Patlar and Civetta 2022).

Interestingly, we also found an underrepresentation of relaxed selected genes on the X-chromosome. In *Drosophila*, as in many other species, males are the hemizygous sex. Hemizygosity allows for recessive mutations to fix on the X chromosome more easily than on the autosomes. For this reason, in males, X-linked alleles will be more exposed to selection, and the effect of both positive and negative selection will be enhanced on them (Charlesworth

et al., 1987). In addition, it was shown that in *Drosophila melanogaster*, male-biased genes were underrepresented on the X-chromosome. Indeed, only 10% of genes on the X chromosome were male-biased, with this percentage increasing to 14-17% in autosomes (Parisi *et al.*, 2003). Because of this, an underrepresentation of relaxed selected genes on the X-chromosome might be expected, as the low number of male-biased genes present on the X-chromosome will be more likely to be under negative or positive selection.

The rapid divergence of SFP genes (Haerty *et al.*, 2007; Wilburn and Swanson 2016; Rowe *et al.*, 2020; Patlar *et al.*, 2021) and their essential roles in postcopulatory processes (Ravi Ram and Wolfner 2007; Avila *et al.*, 2011; Sirot *et al.*, 2014; Schjenken and Robertson 2020; Wigby *et al.*, 2020) has made them prime targets for the study of sexual selection as a potential driver of speciation. Different studies and mathematical models support the idea of forms of intraspecific sexual selection driving diversification and the onset of reproductive barriers and isolation (Kirkpatrick and Ravigne 2002; Gavrilets and Hayashi 2005). By driving different adaptations and phenotypic diversification between populations, sexual selection would lead to the evolution of different reproductive barriers, which would culminate with the establishment of reproductive isolation (Panhuis *et al.*, 2001; Boughman, 2001; Mendelson and Safran 2021). Moreover, it is believed that sexual conflict can increase genetic divergence leading to the onset of reproductive isolation and, therefore, speciation (Rice 1996, 1998; Howard *et al.*, 1998; Parker and Partridge 1998). However, despite of some studies in support of it, the hypothesis of sexual selection being an engine of speciation remains highly controversial (Safran *et al.*, 2013; Williams *et al.*, 2013; Simmons 2018; Decanter *et al.*, 2023; Murali *et al.*, 2023). In fact, some studies suggest that sexual selection is not strong enough by itself to drive a stop in the gene flow, and that natural selection must contribute to the process in synergy with sexual selection (Maan and Seehausen, 2011; Langerhans *et al.*, 2013; Safran *et al.*, 2013). Genes involved in processes of post-mating sexual selection, such as SFP genes,

can provide a source to test sexual selection as a driver of speciation. The finding of an enrichment of positively selected genes that have a reproductive tissue specificity of expression supports the idea that SFP genes can drive different adaptations between species that can lead to the establishment of reproductive isolation. However, our analysis also shows an enrichment of reproductive-tissue specific SFP genes under negative (purifying) selection. This result suggests that the functions of SFP genes in sperm function and competition are essential for reproductive fitness (Ravi Ram and Wolfner 2007; Avila *et al.,* 2011; Sirot *et al.*, 2014; Schjenken and Robertson 2020; Wigby *et al.*, 2020), and are, for this reason, subjected to purifying selection. Interestingly, a previous study that aimed at finding a genetic commonality between sexual selection and speciation focused on three genes (*Acp36DE, CG9997, SP*), and performed intra and interspecific mating experiments to visualize the effects of gene disruption on both CSP and ISC (Castillo and Moyle, 2014). From their results it emerged that both *Acp36DE* and *CG9997*, when disrupted, affected the male's competitive ability both intra and interspecifically. From our results, these two genes are under patterns of negative purifying selection, with low divergence and low variation. The mode of evolution of these two genes and the result of the cited study hints at a crucial function of these genes in the sperm competition process, which has been maintained by negative selection. By knocking down a gene that might have essential and conserved roles on reproduction and fitness, a negative phenotypic effect is expected. On the other hand, *SP* was found to affect ISC, but not CSP when disrupted. Our results indicate *SP* to be under patterns of relaxed selection, which might serve as an explanation of why gene perturbation is not affecting the sperm competitiveness at an interspecific level.

We tested the four genes that showed divergence by positive selection, and that were previously tested and proved to have a role in ISC for roles on CSP. However, the results of our mating experiments did not reveal any effect of these genes on CSP. Heterospecific crosses

between *D. melanogaster* and *D. simulans* produce completely sterile female hybrids while males are lethal (Barbash, 2010). For this reason, the fitness cost of these matings is extremely high, and CSP might be needed to reduce this cost by limiting the number of inviable/sterile offspring (Noriyuki *et al.*, 2012; Leigh *et al.*, 2020). Given the role of female's genotype and the effect of interactions between female's and male's genotypes on the outcomes of sperm competition (Clark and Begun, 1998; Clark *et al.*, 1999; Chow *et al.*, 2010; Giardina *et al.*, 2011; Lüpold *et al.*, 2013; Reinhart *et al.*, 2015) , it is possible that females are sensitive to the negative effects of an heterospecific sperm. This female effect could be mediated by neural genes that have been shown to be important for sperm competition outcomes in *D. melanogaster*, and even in CSP involving *D. simulans* (Chen *et al.*, 2019; Chen *et al.*, 2022). In fact, previous studies have reported that specific female neurons (Tdc2), when inhibited, can cause females to increase the production of first-male progeny, due to the suppression of second-male sperm usage (Chen *et al.*, 2022; Chen *et al*., 2019). A recent proteome analysis conducted on the female's reproductive tract have also suggested that the secretion of species-specific proteins might serve as the molecular mechanism by which females mediate conspecific sperm precedence (McDonough-Goldstein *et al.*, 2021). Knowing the great fitness costs that females experience when mating with a heterospecific male, favouring the ejaculate of conspecific males, even when lacking important proteins for ISC, might be a defensive mechanism to avoid the maladaptive consequences of heterospecific mating.

The fact that the knockdown of these four genes, individually, does not affect CSP seem to be proving a disconnection between the effects of gene manipulation at the intraspecific level (ISC) and heterospecific level (CSP). This might be seen as evidence that post-copulatory sexual selection on SFP genes cannot drive the onset of reproductive isolation between species, as the two processes would not share a common genetic basis. ISC and CSP share some mechanisms (Manier *et al.*, 2013), and conserved processes might be maintained by the force

of negative selection, while species-specific adaptations might be due to shifts of a gene's identity. On the other hand, if the same genes mediate both processes, the disconnection between single genes effects on ISC and CSP might be due to differences in the genetic basis of the two mechanisms. For example, the genetic basis of species differentiation has been proposed to be more polygenic (Mather, 1943) than the genetic architecture of within species variation. Moreover, several studies have highlighted the importance of gene interactions in the establishment of reproductive barriers, further supporting that speciation has a polygenic basis (Dobzhansky 1937; Muller 1942; Perez and Wu 1995; Johnson 2000; Orr and Irving 2001; Tao *et al.*, 2003; Sawamura *et al.*, 2004; Chang and Noor 2007; Tang and Presgraves 2009; Chang *et al.*, 2010; Phadnis 2011). For this reason, reproductive barriers might only be disrupted by the simultaneous manipulation of multiple genes.

6.0 Policy implications

Over the last decades, one of the most used class of pesticides is the one of the Neonicotinoids (Craddock *et al.*, 2019). This category of insecticides has been used for different functions, from crop protections to flea and tick prevention on dogs and cats (Craddock *et al.,* 2019). Neonicotinoids are chemically related to nicotine, and while being much less toxic to humans and mammals compared to old classes of pesticides, a big concern for the use of this substances regards bee populations. As they persist in the environment, non-target organisms, such as bees and other beneficial insects, are exposed to nicotinoids. In laboratory conditions, the exposure to these substances to bee populations resulted in a significant growth reduction and a lower production of new queens (Whitehorn *et al*., 2012). For this reason, the development of non-chemical pesticides should be considered as an essential objective for future research. To avoid the negative effects that chemical-natured pesticides have on the environment, the use of genetically modified organisms (GMO) is becoming a powerful resolution to this problem. GMOs are advantageous in several ways, as they, for example, reduce the accumulation of spray pesticides on crops.

To control the spread and the reproduction of pests, fitness must be reduced in these species of insects. Male's reproductive genes, such as those that encode for seminal fluid proteins, can affect the male's reproductive fitness. The genetic perturbation of genes that are known to affect sperm competition, for example *via* RNA-interference, can have negative outcomes on sperm competition, reducing the male's ability to fertilize the ova. The four genes I focused on in my mating experiments are known to affect intraspecific sperm competition, and they affect this process in a negative way when they are knocked down, lowering the male's fitness. Importantly, I have established here that the perturbation of expression of these genes does not affect interspecific barriers, which suggests that gene manipulations might not spread to other species. The negative consequences of the perturbation of these genes might be

particularly useful in the field of pest control. Specifically, the use of RNA-interference or CRISPR-CAS on sperm competition genes might find a usage in the only pest species of *Drosophila*, *D. suzukii*.

References

Alcock, J. (1994). *Postinsemination Associations Between Males and Females in Insects: The Mate-Guarding Hypothesis*. Annual Review of Entomology, 39(1): 1-21

Ambrosi, P., Chahda, S., Yang, E., Sousa-Neves, R., Mizutani, C.M. (2013). An Evo-Devo Perspective on Hybrid Infertility and Speciation. In: Pontarotti, P. (eds) Evolutionary Biology: Exobiology and Evolutionary Mechanisms. Springer, Berlin, Heidelberg.

Artieri, C. G., Haerty, W., and Singh, R. S. (2007). Association between levels of coding sequence divergence and gene misregulation in Drosophila male hybrids. *Journal of Molecular Evolution*, 65: 697–704

Assis, R., Zhou, Q., and Bachtrog, D. (2012). Sex-biased transcriptome evolution in Drosophila. *Genome Biology and Evolution*, 4(11): 1189-200

Avila, F. W., Ravi Ram, K., Qazi, M. C. B., and Wolfner, M. F. (2010). Sex peptide is required for the efficient release of stored sperm in mated Drosophila females. *Genetics*, 186: 595 – 600

Avila, F. W., Sirot, L. K., LaFlamme, B. A., Rubinstein, C. D., and Wolfner, M. F. (2011). Insect seminal fluid proteins: identification and function. *Annual Review of Entomology*, 56: 21-40

Avila, F.W., Wolfner, M.F. (2009). Acp36DE is required for uterine conformational changes in mated Drosophila females. *Proc. Natl. Acad. Sci. U.S.A.*, 106(37): 15796-15800

Barbash, D. A. (2010). Ninety years of Drosophila melanogaster hybrids. *Genetics*, 186(1): 1-8

Barron, A. B. (2000). Anaesthetising Drosophila for behavioural studies. *J Insect Physiol.*, 46(4): 439-442

Begun, D. J., Lindfors, H. A., Kern, A. D., and Jones, C. D. (2007). Evidence for de novo evolution of testis-expressed genes in the Drosophila yakuba/Drosophila erecta clade. *Genetics*, 176(2):1131-7

Begun, D. J., Whitley, P., Todd, B. L., Waldrip-Dail, H. M., and Clark, A. G. (2000). Molecular Population Genetics of Male Accessory Gland Proteins in Drosophila. *Genetics*, 156(4):1879–1888

Bertram, M. J., Neubaum, D. M., and Wolfner, M. F. (1996). Localization of the Drosophila male accessory gland protein Acp36DE in the mated female suggests a role in sperm storage. *Insect Biochem. Mol. Biol*., 26: 971–980

Birkhead, T. R., and Biggins, J. D. (1998). Sperm competition mechanisms in birds: models and data, *Behavioral Ecology*, 9(3): 253–260

Boellstroff, D. E., Owings, D. H., Penedo, M. C. T., and Hersek, M. J. (1994). Reproducitve behaviour and multiple paternity of California ground squirrels. *Anim. Behav.*, 47: 1057-1064

Bono, J. M., Matzkin, L. M., Hoang, K., and Brandsmeier, L. (2015). Molecular evolution of candidate genes involved in post-mating-prezygotic reproductive isolation. *J Evol Biol*., 28(2):403-14

Boughman, J. W. (2001). Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature*, 411(6840): 944-8

Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csárdi, G., Harrigan, P., Weier, M., Liechti, A., Aximu-Petri, A., Kircher, M., Albert, F. W., Zeller, U., Khaitovich, P., Grützner, F., Bergmann, S., Nielsen, R., Pääbo, S., and Kaessmann, H. (2011). The evolution of gene expression levels in mammalian organs. *Nature*, 78(7369): 343-8

Carroll, S. P. (1991). The adaptive signif- icance of mate guarding in the soap- berry bug, Jadera haematoloma (Hemiptera: Rhopalidae). *J. Insect Behav*., 4: 509-30

Carvalho, G. B., Kapahi, P., Anderson, D. J., and Benzer, S. (2006). Allocrine modulation of feeding behavior by the Sex Peptide of Drosophila. *Curr Biol*., 16(7):692-6

Castillo, D. M., and Moyle, L. C. (2014). Intraspecific sperm competition genes enforce post-mating species barriers in Drosophila. *Proceedings of the Royal Society B*, 281(1797), 2014-2050

Chang, A. S., and Noor, M. A. (2007). The genetics of hybrid male sterility between the allopatric species pair Drosophila persimilis and D. pseudoobscura bogotana: dominant sterility alleles in collinear autosomal regions. *Genetics*, 176(1): 343-9

Chang, A. S., Bennett, S. M., and Noor, M. A. (2010). Epistasis among Drosophila persimilis factors conferring hybrid male sterility with D. pseudoobscura bogotana. *PLoS One*, 5(10): e15377

Chapman, T., Liddle, L., Kalb, J., Wolfner, M. F., and Partridge, L. (1995). Cost of mating in Drosophila melanogaster females is mediated by male accessory gland products. *Nature*, 373: 241–244

Chapman, T., Neubaum, D. M., Wolfner, M. F., and Partridge, L. (2000). The role of male accessory gland protein Acp36DE in sperm competition in Drosophila melanogaster. *Proc Biol Sci.*, 267(1448): 1097-105

Chapman. T., Bangham, J., Vinti, G., Seifried, B., Lung, O., Wolfner, M. F., Smith, H. K., and Partridge, L. (2003). The sex peptide of Drosophila melanogaster: female post-mating responses analyzed by using RNA interference. *Proc Natl Acad Sci U S A*, 100(17): 9923-8

Charlesworth, B., J. A. Coyne, and N. H. Barton. (1987). The Relative Rates of Evolution of Sex Chromosomes and Autosomes. *The American Naturalist*, 130(1): 113–46

Chen, D. S., Clark, A. G., and Wolfner, M. F. (2022). Octopaminergic/tyraminergic *Tdc2* neurons regulate biased sperm usage in female *Drosophila melanogaster*. *Genetics*, 221(4): iyac096

Chen, D. S., Delbare, S. Y. N., White, S. L., Sitnik, J., Chatterjee, M., DoBell, E., Weiss, O., Clark, A. G., and Wolfner, M. F. (2019). Female Genetic Contributions to Sperm Competition in *Drosophila melanogaster*. *Genetics*, 212(3): 789-800

Chen, P. S., Stumm-Zollinger, E., Aigaki, T., Balmer, J., Bienz, M., and Böhlen, P. (1988). A male accessory gland peptide that regulates reproductive behavior of female D. melanogaster. *Cell*, 54(3): 291-298

Chen, P.S., and Bühler, R. (1970). Paragonial substance (sex peptide) and other free ninhydrin-positive components in male and female adults of Drosophila melanogaster. *Journal of Insect Physiology*, 16(4): 615-627

Chow, C. Y., Wolfner, M. F., and Clark, A. G. (2010). The genetic basis for male x female interactions underlying variation in reproductive phenotypes of Drosophila. *Genetics*, 186(4): 1355-65

Civetta A. (1999). Direct visualization of sperm competition and sperm storage in Drosophila. *Curr Biol.*, 9(15): 841-4

Civetta, A., and Clark, A. G. (2000). Correlated effects of sperm competition and postmating female mortality. *Proc Natl Acad Sci U S A*, 97(24):13162-5

Civetta, A., and Ranz, J. M. (2019). Genetic Factors Influencing Sperm Competition. *Front Genet.*, 10:820

Civetta, A., and Singh, R. S. (1998). Sex-related genes, directional sexual selection, and speciation. *Mol Biol Evol*., 15(7): 901-9

Clark, A. G., and Begun, D. J. (1998). Female genotypes affect sperm displacement in Drosophila. *Genetics*, 149(3): 1487-93

Clark, A. G., Begun, D. J., and Prout, T. (1999). Female x male interactions in Drosophila sperm competition. *Science*, 283(5399): 217-20

Clark, A. G., M. Aguade, T. Prout, L. G. Harshman, and C. H. Langley. (1995). Variation in sperm displacement and its association with accessory-gland protein loci in *Drosophila-Melanogaster*. *Genetics*, 139: 189–201

Coyne, J. A., Orr, H. A. (1989). Patterns Of Speciation In Drosophila. *Evolution*, 43(2-1): 362–381

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

Craddock, H.A., Huang, D., Turner, P. C., Quirós-Alcalá, L., and Payne-Sturges, D. C. (2019). Trends in neonicotinoid pesticide residues in food and water in the United States, 1999-2015. *Environ Health*, 18(1):7

Cridland, J. M., Majane, A. C., Sheehy, H. K., and Begun, D. J. (2020). Polymorphism and Divergence of Novel Gene Expression Patterns in *Drosophila melanogaster*. *Genetics*, 216(1): 79-93

Cruzan, M.B. and Arnold, M.L. (1993), Ecological And Genetic Associations In An Iris Hybrid Zone. *Evolution*, 47: 1432-1445

Dapper, A. L., and Wade, M. J. (2020). Relaxed Selection and the Rapid Evolution of Reproductive Genes. *Trends in Genetics*, 36(9): 640-649

Davis, H., Sosulski, N., and Civetta, A. (2020). Reproductive isolation caused by azoospermia in sterile male hybrids of *Drosophila*. *Ecol Evol*, 10(12): 5922-5931

Decanter, N., Normand, R., Souissi, A., Labbé, C., Edeline, E., and Evanno, G. (2023). Sperm competition experiments reveal low prezygotic postmating isolation between parasitic and nonparasitic lamprey ecotypes. *Evolutionary Ecology*, 13(4): e9970

Dobzhansky, T. (1937). Genetics and the origin of species. *Columbia Univ. Press*, New York, NY

Dobzhansky, T. (1950). Evolution in the Tropics. *American Scientist, 38: 209-22*

Dosselli, R., Grassl, J., Den Boer, S. P. A., Kratz, M., Moran, J. M., Boomsma, J. J., *et al.* (2019). Protein-level interactions as mediators of sexual conflict in ants. *Mol. Cell Proteomics*, 18(S34–S45)

E. Mayr. (1942). Systematics and the Origin of Species from the Viewpoint of a Zoologist. *Columbia University Press.*

Ellegren, H., and Parsch, J. (2007). The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet*., 8(9):689-98

Fedorka, K. M., Winterhalter, W. E., and Ware B. (2011). Perceived sperm competition intensity influences seminal fluid protein production prior to courtship and mating. *Evolution*, 65(2): 584-90

Findlay, G. D., Sitnik, J. L., Aquadro, C. F., Clark, N. L., and Wolfner, M. F. (2014). Evolutionary rate covariation identifies new members of a protein network required for Drosophila female post-mating responses. *PLoS Genet.,* 10: e1004108

Findlay, G. D., Yi, X., Maccoss, M. J., and Swanson, W. J. (2008). Proteomics reveals novel Drosophila seminal fluid proteins transferred at mating. *PLoS Biol*., 6(7): e178

Firman, R. C., and Simmons, L. W. (2008). The frequency of multiple paternity predicts variation in testes size among island populations of house mice**.** *J. Evol. Biol.,* 21: 1524-1533

Fisher, H., Jacobs-Palmer, E., Lassance, JM. *et al.* (2016). The genetic basis and fitness consequences of sperm midpiece size in deer mice. *Nat Commun,* **7**: 136-52

Fiumera, A. C. (2007). Associations between sperm competition and natural variation in male reproductive genes on the third chromosome of *Drosophila melanogaster*. *Genetics,* 176: 1245–1260

Fiumera, A. C., B. L. Dumont, and A. G. Clark. (2005). Sperm competitive ability in *Drosophila melanogaster* associated with variation in male reproductive proteins. *Genetics*, 169: 243–257

Flexner, S. B. and Hauck, L. C. (1993). Random House Unabridged Dictionary. 2nd ed., newly rev. and updated. New York, Random House

Freeman, B. G., Rolland, J., Montgomery, G. A., Schluter, D. (2022). Faster evolution of a premating reproductive barrier is not associated with faster speciation rates in New World passerine birds. *Proc. R. Soc. B*, 289: 20211514

Garlovsky, M. D., and Snook, R. R. (2018). Persistent postmating, prezygotic reproductive isolation between populations. *Ecol Evol*, 8(17): 9062-9073

Garlovsky, M. D., Evans, C., Rosenow, M. A., Karr, T. L., and Snook, R. R. (2020). Seminal fluid protein divergence among populations exhibiting postmating prezygotic reproductive isolation. *Mol Ecol*., 29(22): 4428-4441

Gavrilets, S., Hayashi, T. I. (2005). Speciation and sexual conflict. *Evolutionary Ecology*, 19: 167–198

Geyer, L.B. and Palumbi, S.R. (2005), Conspecific Sperm Precedence in two Species of Tropical Sea Urchins. *Evolution*, 59: 97-105

Giardina, T. J., Beavis, A., Clark, A. G., and Fiumera, A. C. (2011). Female influence on pre- and post-copulatory sexual selection and its genetic basis in Drosophila melanogaster. *Molecular Ecology*, 20(19): 4098-108

Go, A. C., and Civetta, A. (2020). Hybrid Incompatibilities and Transgressive Gene Expression Between Two Closely Related Subspecies of *Drosophila*. *Frontiers in Genetics*, 11: 599292

Grafen, A., and Ridley, M. (1983). A model of mate guarding. *Journal of Theoretical Biology*, 102(4)

Greenspan, L., and Clark, A. G. (2011). Associations between Variation in X Chromosome Male Reproductive Genes and Sperm Competitive Ability in Drosophila melanogaster. *Int J Evol Biol.*, 2011:214-280

Grewal, G. (2021). The role of candidate genes within a mapped genome region on sperm competition. Department of biology, University of Winnipeg.

Haerty, W., Jagadeeshan, S., Kulathinal, R. J., Wong, A., Ravi Ram, K., Sirot, L. K., Levesque, L., Artieri, C. G., Wolfner, M. F., Civetta, A., and Singh, R. S. (2007). Evolution in the fast lane: rapidly evolving sex-related genes in Drosophila. *Genetics*, 177(3): 1321-35

Hahn, M. W., and Kern, A. D. (2005). Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks. *Molecular Biology and Evolution*, 22(4): 803-6

Hahn, M. W., Conant, G. C., and Wagner, A. (2004). Molecular evolution in large genetic networks: does connectivity equal constraint? *Journal of Molecular Evolution*, 58(2): 203-11

Harshaman, L. G. and Prout, T. (1994). Sperm displacement without sperm transfer in Drosophila melanogaster. *Evolution*, 48: 758-766

Haygood, R., Babbitt, C. C., Fedrigo, O., and Wray, G. A. (2010). Contrasts between adaptive coding and noncoding changes during human evolution. *Proc Natl Acad Sci U S A*, 107(17):7853-7

Helsen, J., Voordeckers, K., Vanderwaeren, L., Santermans, T., Tsontaki, M., Verstrepen, K. J., and Jelier, R. (2020). Gene Loss Predictably Drives Evolutionary Adaptation. *Mol Biol Evol.*, 37(10): 2989-3002

Hodgins, K. A., Yeaman, S., Nurkowski, K. A., Rieseberg, L. H., and Aitken, S. N. (2016). Expression Divergence Is Correlated with Sequence Evolution but Not Positive Selection in Conifers. *Molecular Biology and Evolution*, 33(6): 1502-16

Holloway, A. K., and Begun, D. J. (2004). Molecular evolution and population genetics of duplicated accessory gland protein genes in Drosophila. *Mol Biol Evol.*, 21(9):1625-8

Holman, L. (2009). Drosophila Melanogaster Seminal Fluid Can Protect the Sperm of Other Males. *Functional Ecology*, 23(1): 180–86

Hopkins, B. R., Sepil, I., and Wigby, S. (2017). Seminal fluid. *Curr Biol.*, 27(11):R404-R405

Hopkins, B. R., Sepil, I., Thézénas, M. L., Craig, J. F., Miller, T., Charles, P. D., Fischer, R., Kessler, B. M., Bretman, A., Pizzari, T., and Wigby, S. (2019). Divergent allocation of sperm and the seminal proteome along a competition gradient in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A*, 116(36):17925-17933

Howard, D. J. (1999). Conspecific Sperm and Pollen Precedence and Speciation. Annual *Review of Ecology and Systematics*, 30: 109-132

Howard, D. J., M. Reece, P. G. Gregory, J. Chu, and M. L. Cain. (1998). The evolution of barriers to fertilization between closely related organisms. *Oxford University Press,New York*, pp. 279–288 in D.J. Howard and S. H. Berlocher, eds. Endless forms: species and speciation

Howard, D. J., Marshall, J. L., Hampton, D. D., Britch, S. C., Draney, M. L., Chu, J., and Cantrell. R. G. (2002). The Genetics of Reproductive Isolation: A Retrospective and Prospective Look with Comments on Ground Crickets. The *American Naturalist*, 159(S3, S8-S21)

Huber, C. D., Kim, B. Y., Marsden, C. D., and Lohmueller, K. E. (2017). Determining the factors driving selective effects of new nonsynonymous mutations. *Proceedings of the National Academy of Sciences*, 114(17): 4465-4470

Hughes, K. A. (1997). Quantitative genetics of sperm precedence in Drosophila melanogaster. *Genetics*, 145(1):139-51

Hughes, K. A., and Leips, J. (2006). Quantitative trait locus analysis of male mating success and sperm competition in Drosophila melanogaster. *Evolution*, 60(7): 1427-34

Hunt, B. G., Ometto, L., Keller, L., and Goodisman, M. A. D. (2013). Evolution at two levels in fire ants: the relationship between patterns of gene expression and protein sequence evolution. *Molecular Biology and Evolution*, 30: 263–271

Hurtado, J., Almeida, F.C., Belliard, S.A., Revale, S. and Hasson, E. (2022) On how to identify a seminal fluid protein: A response to Wigby et al. *Insect Molecular Biology*, 31(5), 537– 542

Iwata, Y., Sato, N., Hirohashi, N., Watanabe, Y., Sauer, W. H. H., and Shaw, P. W. (2021). Sperm competition risk affects ejaculate strategy in terms of sperm number but not sperm size in squid. *J Evol Biol*., 34(9):1352-1361

Jiggins, C., Naisbit, R., Coe, R. and Mallet, J. (2001). Reproductive isolation caused by colour pattern mimicry. *Nature*, 411: 302–305

Johnson, N.A. (2000). Gene interactions and the origin of species. In Epistasis and the Evolutionary Process. Edited by J. Wolf, E. Brodie, and M. Wade. *Oxford University Publishes*, New York. pp. 197–212

Jordan, I. K., Wolf, Y. I., and Koonin, E. V. (2003). No simple dependence between protein evolution rate and the number of protein-protein interactions: only the most prolific interactors tend to evolve slowly. *BMC Ecology and Evolution*, 3: 1

Josephs, E. B., Wright, S. I., Stinchcombe, J. R., and Schoen, D. J. (2017). The Relationship between Selection, Network Connectivity, and Regulatory Variation within a Population of Capsella grandiflora. *Genome Biology and Evolution*, 9(4): 1099-1109

Kawano, N., Araki, N., Yoshida, K., Hibino, T., Ohnami, N., Makino, M., Kanai, S., Hasuwa, H., Yoshida, M., Miyado, K., and Umezawa, A. (2014). Seminal vesicle protein SVS2 is required for sperm survival in the uterus. *Proc Natl Acad Sci U S A*, 111(11): 4145-50

Kawecki, T. J., Barton, N. H., and Fry, J. D. (1997). Mutational collapse of fitness in marginal habitats and the evolution of ecological specialisation. *J Evol Biol.*, 10:407–429

Khaitovich, P., Hellmann, I., Enard, W., Nowick, K., Leinweber, M., Franz, H., Weiss, G., Lachmann, M., and Pääbo, S. (2005). Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science*, 309(5742): 1850-4

Khodursky, S., Svetec, N., Durkin, S. M., and Zhao, L. (2020). The evolution of sex-biased gene expression in the *Drosophila* brain. *Genome Research*, 30(6): 874-884

Kim, D., Langmead, B. and Salzberg, S. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 12, 357–360

Kirkpatrick, M., Ravigne, V. (2002). Speciation by natural and sexual selection: models and experiments. *The American Naturalist*, 159: S22 – S35

Kreitman, M. (2000). Methods to Detect Selection in Populations with Applications to the Human. *Annual Review of Genomics and Human Genetics*, 1(1): 539-559

LaFlamme, B. A., Ram, K. R., and Wolfner, M. F. (2012). The Drosophila melanogaster seminal fluid protease "seminase" regulates proteolytic and post-mating reproductive processes. *PLoS Genet*., 8(1):e1002435

LaMunyon, C. W., and Ward, S. (1998). Larger sperm outcompete smaller sperm in the nematode Caenorhabditis elegans. *Proc. R. Soc. Lond. Ser. B Biol. Sci.*, 26: 1997–2002

Lande, R. (1981). Models of speciation by sexual selection on polygenic traits. *Proc Natl Acad Sci U S A*, 78(6): 3721-5

Langerhans, R. B., and Riesch, R. (2013). Speciation by selection: A framework for understanding ecology's role in speciation, *Current Zoology*, 59(1): 31–52

Lawniczak, M. K., and Begun, D. J. (2005). A QTL analysis of female variation contributing to refractoriness and sperm competition in Drosophila melanogaster. *Genet Res*., 86(2): 107-14

Leiblich, A., Marsden, L., Gandy, C., Corrigan, L., Jenkins, R., Hamdy, F., and Wilson, C. (2012). Bone morphogenetic protein- and mating-dependent secretory cell growth and migration in the Drosophila accessory gland. *Proc Natl Acad Sci U S A*, 109(47):19292-7

Leigh, S., Rostant, W. G., Taylor, M. I., Alphey, L., and Chapman, T. (2021). Satyrization in Drosophila fruitflies. *J Evol Biol*., 34(2): 319-330

Lemos, B., Bettencourt, B. R., Meiklejohn, C. D., and Hartl, D. L. (2005). Evolution of proteins and gene expression levels are coupled in Drosophila and are independently associated with

mRNA abundance, protein length, and number of protein-protein interactions. *Molecular Biology and Evolution*, 22(5): 1345-54

Levine, M. T., Jones, C. D., Kern, A. D., Lindfors, H. A., and Begun, D. J. (2006). Novel genes derived from noncoding DNA in Drosophila melanogaster are frequently X-linked and exhibit testis-biased expression. *Proc Natl Acad Sci U S A*, 103(26):9935-9

Liao, B. Y., Weng, M. P, and Zhang, J. (2010). Contrasting genetic paths to morphological and physiological evolution. *Proc Natl Acad Sci U S A*, 107(16): 7353-8

Liberti, J., Baer, B. and Boomsma, J.J. (2018). Rival seminal fluid induces enhanced sperm motility in a polyandrous ant. *BMC Evol Biol*, 18: 28

Liu, H., and Kubli, E. (2003). Sex-peptide is the molecular basis of the sperm effect in Drosophila melanogaster. *Proc Natl Acad Sci U S A*, 100(17):9929-33

Lung, O., and Wolfner, M.F. (1999). Drosophila seminal fluid proteins enter the circulatory system of the mated female fly by crossing the posterior vaginal wall. *Insect Biochemistry and Molecular Biology*, 29(12): 1043-1052

Lüpold, S., Pitnick, S., Berben, K. S., Blengini, C. S., Belote, J. M., and Manier, M. K. (2013). Female mediation of competitive fertilization success in Drosophila melanogaster. *Proceedings of the National Academy of Sciences*, 110(26): 10693-8

Maan, M.E. and Seehausen, O. (2011), Ecology, sexual selection and speciation. *Ecology Letters*, 14: 591-602

Manier, M. K., Belote, J. M., Berben, K. S., Lüpold, S., Ala-Honkola, O., Collins, W. F., and Pitnick, S. (2013). Rapid diversification of sperm precedence traits and processes among three sibling Drosophila species. *Evolution*, 67(8): 2348-62

Manier, M. K., Belote, J. M., Berben, K. S., Novikov, D., Stuart, W. T., and Pitnick, S. (2010). Resolving mechanisms of competitive fertilization success in Drosophila melanogaster. *Science*, 328(5976): 354-7

Manier, M. K., Lüpold, S., Belote, J. M., Starmer, W. T., Berben, K. S., Ala-Honkola, O., Collins, W. F., and Pitnick, S. (2013). Postcopulatory sexual selection generates speciation phenotypes in Drosophila. *Current Biology*, 23(19): 1853-62

Mank, J. E., Hultin-Rosenberg, L., Zwahlen, M., and Ellegren, H. (2008). Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. *Am Nat.*, 171(1):35-43

Manning, A. (1962). A Sperm Factor Affecting the Receptivity of *Drosophila Melanogaster* Females. *Nature*, 194: 252–253

Manning, A. (1967). The control of sexual receptivity in female Drosophila. *Anim Behav.*, 15(2):239-50

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

Mather, K. (1943), Polygenic Inheritance And Natural Selection. *Biological Reviews*, **18**: 32-64

McDonough-Goldstein, C. E., Whittington, E., McCullough, E. L., Buel, S. M., Erdman, S., Pitnick, S., and Dorus, S. (2021). Pronounced Postmating Response in the Drosophila Female Reproductive Tract Fluid Proteome. *Molecular & Cellular Proteomics*, 20: 100156

McGraw, L. A., Gibson, G., Clark, A. G., and Wolfner, M. F. (2004). Genes regulated by mating, sperm, or seminal proteins in mated female Drosophila melanogaster. *Curr. Biol*., 14:1509–14

Meiklejohn, C. D., Parsch, J., Ranz, J. M., and Hartl, D. L. (2003) Rapid evolution of male-biased gene expression in Drosophila. *Proceedings of the National Academy of Sciences*, 100(17): 9894-9

Mendelson, T. C., and Safran, R. J. (2021). Speciation by sexual selection: 20 years of progress. *Trends Ecol Evol*., 36(12): 1153-1163

Misra, S., and Wolfner, M. F. (2020). *Drosophila* seminal sex peptide associates with rival as well as own sperm, providing SP function in polyandrous females. *Elife*, 9:e58322

Monsma, S.A., Harada, H.A., and Wolfner, M.F. (1990). Synthesis of two Drosophila male accessory gland proteins and their fate after transfer to the female during mating. *Dev. Biol.*, 142: 465-475

Mozuraitis, R., Murtazina, R., Zurita, J. Pei, Y., Ilag, L., Wiklund, C., Karlson, A. K. B. (2019). Anti-aphrodisiac pheromone, a renewable signal in adult butterflies. *Sci Rep,* 9: 142-62

Mueller, J. L., Page, J. L., and Wolfner, M. F. (2006). An ectopic expression screen reveals the protective and toxic effects of Drosophila seminal fluid proteins. *Genetics*, 175(2):777-83

Mueller, J. L., Ravi Ram, K., McGraw, L. A., Blochn Qazi, M. C., Siggia, E. D., Clark, A. G., Aquadro, C. F., and Wolfner, M. F. (2005). Cross-species comparison of Drosophila male accessory gland protein genes. *Genetics,* 171: 131–143

Mueller, J.L., Linklater, J.R., Ravi Ram, K., Chapman, T., Wolfner, M.F. (2008). Targeted gene deletion and phenotypic analysis of the Drosophila melanogaster seminal fluid protease inhibitor Acp62F. *Genetics,* 178(3): 1605-14

Muller, H. J. (1942). Isolating mechanisms, evolution and temperature. *Biological Symposia*, 6: 71–125. Lancaster, Penn.: Jaques Cattell Press.

Murali, G., Meiri, S., and Roll, U. (2023). Chemical signalling glands are unlinked to species diversification in lizards. *Evolution*, qpad101

Nakadera, Y., Thornton Smith, A., Daupagne, L., Coutellec, M. A., Koene, J. M., and Ramm, S. A. (2020). Divergence of seminal fluid gene expression and function among natural snail populations. *J Evol Biol*., 33(10):1440-1451

Neubaum, D. M., and Wolfner, M. F. (1999). Mated Drosophila melanogaster females require a seminal fluid protein, Acp36DE, to store sperm efficiently. *Genetics*, 153(2): 845-57

Neubaum, D. M., and Wolfner, M. F. (1999). Wise, winsome, or weird? Mechanisms of sperm storage in female animals. *Curr Top Dev Biol*., 41: 67–97

Nonidez, J. F. (1920). The Phenomena of internal reproduction in Drosophila. *Biol. Bull,* 39: 208–30

Noriyuki, S., Osawa, N. and Nishida, T. (2012), Asymmetric reproductive interference between specialist and generalist predatory ladybirds. *Journal of Animal Ecology*, 81: 1077-1085

Nosil, P. (2012). *Ecological Speciation*, *Oxford Series in Ecology and Evolution*. Oxford.

Nuzhdin, S. V., Wayne, M. L., Harmon, K. L., and McIntyre, L. M. (2004). Common pattern of evolution of gene expression level and protein sequence in Drosophila. *Molecular Biology and Evolution*, 21(7): 1308-17

Orr, H. A. (2005). The genetic basis of reproductive isolation: insights from Drosophila. Proc Natl Acad Sci U S A, 102(Suppl 1): 6522-6

Orr, H. A., and Irving, S. (2001). Complex epistasis and the genetic basis of hybrid sterility in the Drosophila pseudoobscura Bogota-USA hybridization. *Genetics*, 158(3): 1089-100

Palmieri, N., Kosiol, C., and Schlötterer, C. (2014). The life cycle of Drosophila orphan genes. *Elife*, 3:e01311

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics*, 25(1): 547-572

Panhuis, T. M., Butlin, R., Zuk, M., and Tregenza, T. (2001). Sexual selection and speciation. *Trends Ecol Evol*., 16(7): 364-371

Papakostas, S., Vøllestad, L. A., Bruneaux, M., Aykanat, T., Vanoverbeke, J., Ning, M., Primmer, C. R., and Leder, E. H. (2014). Gene pleiotropy constrains gene expression changes in fish adapted to different thermal conditions. *Nature Communications*, 5: 4071

Parisi, M., Nuttall, R., Edwards, P., Minor, J., Naiman, D., Lü, J., Doctolero, M., Vainer, M., Chan, C., Malley, J., Eastman, S., and Oliver, B. (2004). A survey of ovary-, testis-, and soma-biased gene expression in Drosophila melanogaster adults. *Genome Biology and Evolution*, 5(6):R40

Parisi, M., Nuttall, R., Naiman, D., Bouffard, G., Malley, J., Andrews, J., Eastman, S., and Oliver, B. (2003). Paucity of genes on the Drosophila X chromosome showing male-biased expression. *Science*, 299(5607): 697-700

Parker, G. A. (1979). Sexual selection and sexual conflict. In Sexual selection and reproductive competition in insects. (ed. Blum MS, Blum NA), pp. 123–166 *Academic, New York*

Parker, G. A. (1990). Sperm competition games: raffles and roles. *Proc. R. Soc. Lond. B.*, 24: 2120–126

Parker, G. A. (1998). Sperm competition and the evolution of ejaculates: Towards a theory base. In T. R. Birkhead & A. P. Møller (Eds.), *Sperm competition and sexual selection* (pp. 3–54). New York: Academic.

Parker, G. A., and Partridge, L. (1998). Sexual conflict and speciation. *Philos Trans R Soc Lond B Biol Sci.*, 353(1366): 261-74

Parker, G.A. (1970). Sperm Competition And Its Evolutionary Consequences In The Insects. *Biological Reviews*, 45: 525-567

Parker, G.A. (1979). Sexual Selection and Reproductive Competition in Insects. *Academic Press*, pp. 123-166

Patlar, B., and Civetta, A. (2022). Seminal fluid gene expression and reproductive fitness in Drosophila melanogaster. *BMC Ecology and Evolution*, 22(1): 20

Patlar, B., Jayaswal, V., Ranz, J. M., and Civetta, A. (2021). Nonadaptive molecular evolution of seminal fluid proteins in Drosophila. *Evolution*, 75(8), 2102-2113

Peng, J., Chen, S., Büsser, S., Liu, H., Honegger, T., and Kubli, E.(2005). Gradual release of sperm bound sex-peptide controls female postmating behavior in Drosophila. *Curr Biol*., 15(3):207-13

Perez, D. E., and Wu, C. I. (1995). Further characterization of the Odysseus locus of hybrid sterility in Drosophila: one gene is not enough. *Genetics*, 140(1): 201-6

Perotti, M. E. (1973). The mitochondrial derivative of the spermatozoon of Drosophila before and after fertilization. *J Ultrastruct Res*., 44(3):181-98

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, 33(3): 290-5

Phadnis, N. (2011). Genetic architecture of male sterility and segregation distortion in Drosophila pseudoobscura Bogota-USA hybrids. *Genetics*, 189(3): 1001-9

Pilpel, N., Nezer, I., Applebaum, S. W., and Heifetz, Y. (2008). Mating-increases trypsin in female Drosophila hemolymph. *Insect Biochem Mol Biol*., 38(3):320-30

Pitnick, S., Markow, T. A., and Spicer, G. S. (1999). Evolution of multiple kinds of female sperm-storage organs in *Drosophila*. *Evolution*, 53: 1804–1822

Polak, M., Wolf, L., Starmer, W. *et al.* (2001). Function of the mating plug in *Drosophila hibisci* Bock. *Behav Ecol Sociobiol.*, 49: 196–205

Preston, B. T., Stevenson, I. R., Pemberton, J. M., Coltman, D. W., and Wilson, K. (2003). Overt and covert competition in a promiscuous mammal: the importance of weaponry and testes size to male reproductive success. *Proc Biol Sci*, 270(1515):633-4

Price, C. (1997). Conspecific sperm precedence in Drosophila. *Nature*, 388: 663–666

Price, C. S., Kim, C. H., Posluszny, J., and Coyne, J. A. (2000). Mechanisms of conspecific sperm precedence in Drosophila. *Evolution*, 54: 2028–2037

Price, C.S.C, Dyer, K.A, and Coyne, J.A. (1999). Sperm competition between Drosophila males involves both displacement and incapaitation. *Nature*, 400: 449–452

Pröschel, M., Zhang, Z., and Parsch, J. (2006). Widespread adaptive evolution of Drosophila genes with sex-biased expression. *Genetics*, 174(2):893-900

Purandare, S. R., Bickel, R. D., Jaquiery, J., Rispe, C., and Brisson, J. A. (2014). Accelerated evolution of morph-biased genes in pea aphids. *Mol Biol Evol*., 31(8): 2073-83

Ram, K.R., Wolfner, M.F. (2007). Sustained post-mating response in Drosophila melanogaster requires multiple seminal fluid proteins. *PLoS Genet*., 3(12):e238

Ramm, S.A., Edward, D.A., Claydon, A.J. *et al.* (2015). Sperm competition risk drives plasticity in seminal fluid composition. *BMC Biol.*, 13: 87

Ranz, J. M., Castillo-Davis, C. I., Meiklejohn, C. D., and Hartl, D. L. (2003). Sex-Dependent Gene Expression and Evolution of the Drosophila Transcriptome. *Science*, 300: 1742-1745

Ravi Ram, K., and Wolfner, M. F. (2007). Seminal influences: Drosophila Acps and the molecular interplay between males and females during reproduction. *Integrative and Comparative Biology*, 47(3): 427-45

Ravi Ram, K., and Wolfner, M. F. (2007). Sustained post-mating response in Drosophila melanogaster requires multiple seminal fluid proteins. *PLoS Genet*., 3: e238

Ravi Ram, K., and Wolfner, M. F. (2009). A network of interactions among seminal proteins underlies the long-term postmating response in Drosophila. *Proc Natl Acad Sci*, 106: 15384–15389

Reinhart, M., Carney, T., Clark, A. G., and Fiumera, A. C. (2015). Characterizing male-female interactions using natural genetic variation in Drosophila melanogaster. *Journal of Heredity*, 106(1): 67-79

Rice, W. R. (1996). Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. *Nature*, 381(6579):232-4

Rice, W. R.(1998). Intergenomic conflict, interlocus antagonisic coevolution, and the evolution of reproductive isolation. *Oxford University Press,New York*, pp. 261-270 in D.J. Howard and S. H. Berlocher, eds. Endless forms: species and speciation.

Robinson, T., Johnson, N. and Wade, M. (1994). Postcopulatory, prezygotic isolation: intraspecific and interspecific sperm precedence in Tribolium spp., flour beetles. *Heredity*, 73: 155–159

Rowe, M., Veerus, L., Trosvik, P., Buckling, A., and Pizzari, T. (2020). The Reproductive Microbiome: An Emerging Driver of Sexual Selection, Sexual Conflict, Mating Systems, and Reproductive Isolation. *Trends in Ecology & Evolution*, 35(3): 220-234

Rowley, A. G., Daly-Engel, T. S., and Fitzpatrick, J. L. (2019). Testes size increases with sperm competition risk and intensity in bony fish and sharks, *Behavioral Ecology*, 30(2): 364–371

Rowley, A., Locatello, L., Kahrl, A., Rego, M., Boussard, A., Garza-Gisholt, E., Kempster, R. M., Collin, S. P., Giacomello, E., Follesa, M. C., Porcu, C., Evans, J. P., Hazin, F., Garcia-Gonzalez, F., Daly-Engel, T., Mazzoldi, C., and Fitzpatrick, J. L. (2019). Sexual selection and the evolution of sperm morphology in sharks. *J Evol Biol*., 32(10): 1027-1035

Safran, R. J. , Scordato, E. S. , Symes, L. B. , Rodríguez, R. L. , and Mendelson, T. C. (2013). Contributions of natural and sexual selection to the evolution of premating reproductive isolation: A research agenda. *Trends in ecology & evolution*, 28: 643–650

Sartor, M. A., Zorn, A. M., Schwanekamp, J. A., Halbleib D., Karyala, S., Howell, M. L., Dean, G. E., Medvedovic, M., and Tomlinson, C. R. (2006). A New Method to Remove Hybridization Bias for Interspecies Comparison of Global Gene Expression Profiles Uncovers an Association

between mRNA Sequence Divergence and Differential Gene Expression in Xenopus. *Nucleic Acids Researc,* 34 (1): 185–200

Satoshi A., Dik H., Hiroyuki M., and Masanori K. (2006). Testis size depends on social status and the presence of male helpers in the cooperatively breeding cichlid *Julidochromis ornatus*. *Behavioral Ecology*, 17(3): 372–379

Saudan, P., Hauck, K., Soller, M., Choffat, Y., Ottiger, M., Spörri, M., Ding, Z., Hess, D., Gehrig, P. M., Klauser, S., Hunziker, P., and Kubli, E. (2002). Ductus ejaculatorius peptide 99B (DUP99B), a novel Drosophila melanogaster sex-peptide pheromone. *Eur J Biochem.*, 269(3): 989-97

Sawamura, K., Roote, J., Wu, C. I., and Yamamoto, M. T. (2004). Genetic complexity underlying hybrid male sterility in Drosophila. *Genetics*, 166(2): 789-96

Schjenken, J. E., and Robertson, S. A. (2020). The female response to seminal fluid. *Physiological Reviews*, 100: 1077–1117

Schluter, Dolph, and Associate Editor: Mark A. McPeek. (2000). Ecological Character Displacement in Adaptive Radiation. *The American Naturalist*, 156(S4: S4–16)

Seeholzer, L. F., Seppo, M., Stern, D. L., Ruta, V. (2018). Evolution of a central neural circuit underlies Drosophila mate preferences. *Nature*, 559(7715): 564-569

Sepil, I., Hopkins, B. R., Dean, R., Thézénas, M. L., Charles, P. D., Konietzny, R., Fischer, R., Kessler, B. M., and Wigby, S. (2019). Quantitative Proteomics Identification of Seminal Fluid Proteins in Male *Drosophila melanogaster*. *Mol Cell Proteomics*, 18(1):S46-S58

Short R. V. (1975). The contribution of the mule to scientific thought. *J Reprod Fertil Suppl*, 23: 359-64

Simmerl, E., Schäfer, M., and Schäfer, U. (1995). Structure and regulation of a gene cluster for male accessory gland transcripts in Drosophila melanogaster. *Insect Biochem Mol Biol*., 25(1): 127-37

Simmons, L. W. (2002). Sperm Competition and Its Evolutionary Consequences in the Insects. *Princeton: Princeton University Press*.

Simmons, L. W. (2018). Sperm competition, sexual conflict, and speciation: A comment on Tinghitella et al. *Behavioral Ecology*, 29: 800

Simmons, L. W., and Beveridge, M. (2011). Seminal Fluid Affects Sperm Viability in a Cricket. *PLOS ONE*, 6(3): e17975

Simmons, L. W., and García-González, F. (2008). Evolutionary Reduction in Testes Size and Competitive Fertilization Success in Response to the Experimental Removal of Sexual Selection in Dung Beetles. *Evolution*, 62(10): 2580–91

Simpson, G. G. (1951). The Species Concept. *Evolution*, 5(4- 1)

Simpson, G. G. (1961). A Basic Work On Evolutionary Theory. *Evolution*, 15(1): 112-113

Sirot, L. K., Findlay, G., Sitnik, J., Frasheri, D., Avila, F., and Wolfner, M. F. (2014). Molecular characterization and evolution of a gene family encoding both female- and male-specific reproductive proteins in *Drosophila*. *Molecular Biology and Evolution*, 31: 1554–1567

Swanson, W. J., and Vacquier, V. D. (2002). The rapid evolution of reproductive proteins. *Nat Rev Genet.*, 3(2): 137-44

Swanson, W. J., Clark, A. G., Waldrip-Dail, H. M., Wolfner, M. F., and Aquadro, C. F. (2001). Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in Drosophila. *Proc Natl Acad Sci U S A*, 98(13): 7375-9

Tang, S., and Presgraves, D. C. (2009) Evolution of the Drosophila nuclear pore complex results in multiple hybrid incompatibilities. *Science*, 323(5915): 779-82

Tao, Y., Zeng, Z. B., Li, J., Hartl, D. L., and Laurie, C. C. (2003). Genetic dissection of hybrid incompatibilities between Drosophila simulans and D. mauritiana. II. Mapping hybrid male sterility loci on the third chromosome. *Genetics*, 164(4): 1399-418

Thornhill, R. (1983). Cryptic female choice and its implications in the scorpionfly *Harpobittacus nigriceps*. *Am Nat*, 122: 765-788

Thurmond, J., Goodman, J. L., Strelets, V. B., Attrill, H., Gramates, L. S., Marygold, S. J., Matthews, B. B., Millburn, G., Antonazzo, G., Trovisco, V., Kaufman, T. C., and Calvi, B. R. FlyBase Consortium. (2019). FlyBase 2.0: the next generation. *Nucleic Acids Research*, 47(D1): D759-D765

Tourmente, M., Gomendio, M. and Roldan, E.R. (2011). Sperm competition and the evolution of sperm design in mammals. *BMC Evol Biol.*, 11: 12

Tram, U., and Wolfner, M. F. (1999). Male seminal fluid proteins are essential for sperm storage in Drosophila melanogaster. *Genetics*, 153(2): 837-44

Turissini, D. A., McGirr, J. A., Patel, S. S., David, J. R., and Matute, D. R. (2018). The Rate of Evolution of Postmating-Prezygotic Reproductive Isolation in Drosophila. *Molecular Biology and Evolution*, 35(2): 312–334

Vallender, E. J. (2009). Bioinformatic approaches to identifying orthologs and assessing evolutionary relationships. *Methods*, 49(1):50-5

Van Dyken, J. D., and Wade, M. J. (2010). The genetic signature of conditional expression. *Genetics*, 184(2): 557-70

Vedanayagam, J. P., and Garrigan, D. (2015). The effects of natural selection across molecular pathways in Drosophila melanogaster. *BMC Ecology and Evolution*, 15: 203

Voolstra, C., Tautz, D., Farbrother, P., Eichinger, L., and Harr, B. (2007). Contrasting evolution of expression differences in the testis between species and subspecies of the house mouse. *Genome Res.*, 17(1):42-9

Waage, J. K. (1986). Evidence for widespread sperm displacement ability among Zygoptera (Odonata) and the means for predicting its existence. *Biological Journal of the Linnean Society,* 28: 285–300

Warnefors, M., and Kaessmann, H. (2013). Evolution of the correlation between expression divergence and protein divergence in mammals. *Genome Biology and Evolution*, 5(7): 1324-35

Whitehorn, P. R., O'Connor, S., Wackersand, F. L., and Goulson, D. (2012). Neonicotinoid Pesticide Reduces Bumble Bee Colony Growth and Queen Production. *Science*, 336: 6079

Wigby, S., and Chapman, T. (2004). Sperm competition. *Curr Biol*., 14(3): R100-2

Wigby, S., and Chapman, T. (2005). Sex peptide causes mating costs in female Drosophila melanogaster. *Curr Biol.*, 15(4):316-21

Wigby, S., Brown, N. C., Allen, S. E., Misra, S., Sitnik, J. L., Sepil, I., Clark, A. G., and Wolfner, M. F. (2020). The *Drosophila* seminal proteome and its role in postcopulatory sexual selection. *Philosophical Transactions of the Royal Society B*, 375(1813): 20200072

Wigby, S., Sirot, L. K., Linklater, J. R., Buehner, N., Calboli, F. C., Bretman, A., Wolfner, M. F., and Chapman, T. (2009). Seminal fluid protein allocation and male reproductive success. *Curr Biol*., 19(9): 751-7

Wilburn, D. B., and Swanson, W. J. (2016). From molecules to mating: Rapid evolution and biochemical studies of reproductive proteins. *Journal of Proteomics*, 135: 12-25

Williams, T. H., Gumm, J. M., and Mendelson, T. C. (2013). Sexual selection acting on a speciation trait in darters (Percidae: *Etheostoma*), *Behavioral Ecology*, 24(6): 1407–1414

Wolfner, M. F. (1997). Tokens of love: functions and regulation of Drosophila male accessory gland products. *Insect Biochem Mol Biol*., 27(3):179-92

Wong, A., Albright, S.N., Giebel, J.D., Ram, K.R., Ji, S., Fiumera, A.C., Wolfner, M.F. (2008). A role for Acp29AB, a predicted seminal fluid lectin, in female sperm storage in Drosophila melanogaster. *Genetics*, 180(2): 921-31

Wray, G. (2007). The evolutionary significance of *cis*-regulatory mutations. *Nat Rev Genet.*, 8: 206–216

Zhang, R., Clark , A. G., and Fiumera, A. C. (2013). Natural genetic variation in male reproductive genes contributes to nontransitivity of sperm competitive ability in Drosophila melanogaster. *Mol Ecol.*, 22(5):1400-15

Zhong, X., Lundberg, M., and Råberg, L. (2021). Divergence in Coding Sequence and Expression of Different Functional Categories of Immune Genes between Two Wild Rodent Species. *Genome Biology and Evolution*, 13(3): evab023

Appendices

Supplementary Data

Figure S1. **Electrophoresis on agarose gel for primer testing.** In picture **A**, *D. melanogaster* bands are displayed on the top part of the gel, while *D. simulans* bands are on the bottom part of the gel. Bands for *Sfp33A2* and *Sfp33A4* are missing in *D. simulans*, while *Acp98AB* is showing faint bands in both the species. In picture **B**, the bands for the two species are side to side for each gene, with the band on the left being the one for *D. melanogaster* and the band on the right being the one for *D. simulans*. *CG8157* is not showing bands for neither of the two species. *Mst57Da* is showing faint bands in *D. simulans*, suggesting a low primer efficiency.

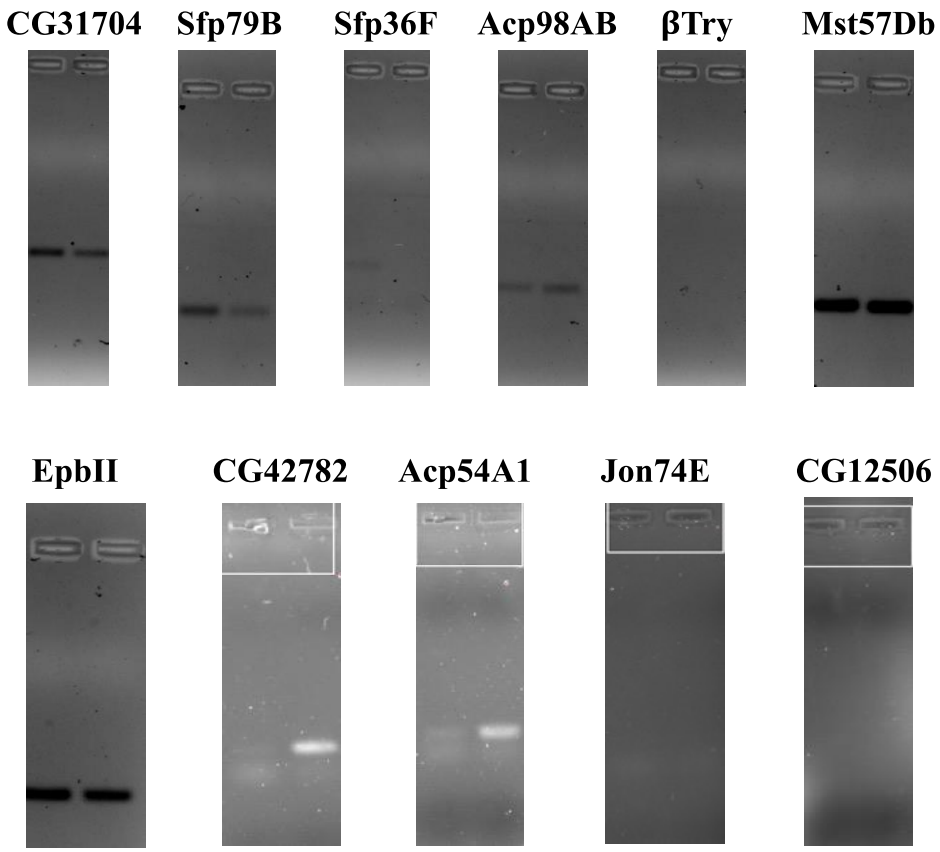

Figure S2. **Electrophoresis on agarose gel with *D. melanogaster* and *D. simulans* cDNA to visualize differences in gene expression.** For each gene, the *D. melanogaster* bands are the ones on the left, while the *D. simulans* bands are the ones on the right. None of these genes is showing a significant lower expression in *D. simulans* compared to *D. melanogaster* (*CG31704*, *Sfp79B*, *Acp98AB*, *Mst57Db*, *EpbII*, *CG42782*, *Acp54A1*) or aren't showing bands for neither of the two species. (*βTry*, *Jon74E*, *CG12506*). *Sfp36F* is not showing a band in *D. simulans*, suggesting a major difference in gene expression. However, amplification of *Sfp36F* in *D. simulans* was poor (**Figure S1**), indicating that this result might only be the consequence of a poor primer annealing.

Supplementary Protocols

1. Homogenization buffer (preparation for 10 mL)

Component	Volume
10M Tris-HCl pH 8.0	100µL
0.5 EDTA	2mL
10% SDS	1mL
Sterile water	6mL

Procedure:

1) Homogenize flies in the homogenization buffer
2) Incubate at 70°C for 30 minutes
3) Add 14µL of 8M potassium acetate and leave on ice for 30 minutes
4) Centrifuge for 20 minutes at 14,000 RPM
5) Add 50mL of 100% isopropanol and let sit at room temperature for 10 minutes
6) Centrifuge for 10 minutes at 14000 RPM
7) Wash twice with 40µL of cold 70% ethanol
8) Let air dry for 40 minutes
9) Suspend it in 40µL of sterile water

2. PCR (Bio Basic Kit)

Preparation for 1x reaction

Component	Volume
Sterile water	13.075 µL
10x Buffer	2.5 µL
MgSO4 (20mM)	3.125 µL
dNTP (10mM)	0.6 µL
Taq polymerase	0.2 µL
DNA template	2 µL
F primer (5mM)	1.0 µL
R primer (5mM)	1.0 µL

PCR program used in Bio-Rad MJ Mini Personal Thermal Cycler is as follows:

1) 95˚C for 5 minutes
2) 95˚C for 1 minute
3) 60˚C for 2 minutes
4) 72˚C for 3 minutes
5) Repeat from step two for a total of 36 times
6) 72˚C for 5 minutes

3. RNA extraction (BioRad Aurum Total RNA Mini Kit)

Component	Volume (µl)
Lysis solution	350 µl

60% ethanol	150 µl
Low stringency wash solution	1400 µl
High stringency wash solution	700 µl
Diluted DNase I	80 µl
Elution solution	40 µl

Procedure:

1) Add 150 µl of lysis solution into Eppendorf tubes
2) Transfer dissected tissues into Eppendorf tubes and add 200 µl more lysis solution
3) Mix the solution by pipetting up and down until froth is showing, as a sign of good mixing
4) Add 350 µl of 60% ethanol
5) Transfer the sample solution into the RNA binding column
6) Centrifuge for 60 sec at 14000 RPM
7) Discard filtrate
8) Add 700 µl of low stringency wash solution to the RNA binding column
9) Centrifuge for 30 sec at 14000 RPM
10) Discard the low stringency wash solution from the wash tube
11) Add 80 µl of diluted DNase I to the membrane stack at the bottom of the column and allow to incubate at room temperature for 25 min
12) Add 700 µl of high stringency wash solution to the RNA binding column
13) Centrifuge for 30 sec
14) Discard the high stringency wash solution from the wash tube
15) Add 700 µl of low stringency was solution to the RNA binding column
16) Centrifuge for 1 min at 14000 RPM
17) Discard the low stringency wash solution from the wash tube
18) Centrifuge for an additional 2 min to remove residual wash solution
19) Transfer the RNA binding column to a 1.5 ml capped microcentrifuge tube
20) Pipette 40 µl of the elution solution onto the membrane stack at the bottom of the RNA binding column and allow 1 min for the solution to saturate the membranes
21) Centrifuge for 2 min at 14000 RPM to elute the total RNA

4. cDNA synthesis (BioRad iScript Select cDNA synthesis)

Component	Volume
5x iScript reaction mix	4 µl
iScript reverse transcriptase	1 µl
RNA template	2 µl
Nuclease-free water	13 µl

cDNA synthesis program used in Bio-Rad MJ Mini Personal Thermal Cycler is as follows:

1) 25°C for 5 min
2) 46°C for 20 min
3) 95°C for 1 min
4) Hold at 10°C
5) End