TRANSCRIPTIONAL DOWN REGULATION OF SPERMATOGENESIS GENES IN *DROSOPHILA* INTERSPECIES HYBRIDS

BY

VIGNESH SUNDARARAJAN

A Thesis submitted to the Faculty of Graduate Studies in Partial Fulfillment of the Requirements for the degree of:

MASTER'S OF SCIENCE

Department of Biology
University of Winnipeg
Winnipeg, Manitoba, Canada

Abstract

In *Drosophila*, crossing two closely related species can generate viable but sterile male offspring, typically an outcome of postzygotic reproductive isolation. Hybrids between species of the *Drosophila simulans* clade show disruption mainly after the meiosis stage of the spermatogenesis pathway, which eventually affects the production of mature sperm. Whole genome investigations (using microarray) of the clade identified that misregulation in sterile hybrids was caused by post-meiotic breakdown. However, either the use of non species-specific genomic platforms or the choice of tissue sampling (whole bodies rather than testes) has made the results of previous investigations prone to ambiguity. In this thesis, I utilized a robust, gene- and species-specific amplification method (quantitative RT-PCR) to analyze the expression levels of spermatogenesis genes from all developmental stages in sterile hybrids and fertile parental species. I found that two of the mitotic genes (*bam* and *bgcn*) and meiotic genes (*can* and *sa*) showed significantly lower expression in the testes of sterile hybrids relative to the parental species. Down regulation of spermatogenesis genes specific to interspecies hybrid testes was further supported by lack of differences in gene expression in hybrid ovaries (interspecific), hybrid whole body samples (interspecific) and hybrid testes (intrapsecific) when compared to their respective parental species. However a preliminary protein assay did not suggest a difference in expression between *D. mauritiana* and interspecies sterile hybrid for *bam* and *sa*. These results suggest that misregulation in hybrid sterile males is solely transcriptional and exclusive to testes in interspecies hybrids. The results presented in this thesis do not support down regulation driven by hybrid male sterility. It is possible that transcriptional down regulation of spermatogenesis genes in interspecific hybrids

could be the result of rapid divergence experienced by the male genome among the closely related species of *Drosophila* (i.e. the male sex drive hypothesis). Alternatively, allometric changes due to subtle testes-specific developmental abnormalities in sterile interspecific hybrids, as suggested by morphological analysis, might also contribute to differences in gene expression.

Acknowledgments

This thesis work would have been impossible without the contribution of a great many people, who I am indebted. I owe my deepest gratitude to my supervisor **Prof. Dr. Alberto Civetta,** for his guidance, enthusiasm, encouragement and support. Dr. Civetta's expertise, understanding and patience, added considerably to my graduate experience as well as the pursuit of my dream to research in molecular genetics. I have been fortunate to have an advisor who gave me the freedom to explore, but still gave timely constructive criticisms. It was my privilege to have **Dr. Jens Franck** as a member in my committee and for the advice he provided at all levels towards the betterment of my project. I am grateful to **Dr. Sara Good-Avila** for insightful comments and suggestions on my thesis as a committee member. I specially thank her for encouraging the use of correct grammar and consistent notation in my writings. I would like to thank **Dr. Steve Whyard** from the University of Manitoba and **Dr. Sandra Kirby** for taking time out from their busy schedules to serve on my thesis examination committee.

My special thanks always goes out to **Barb Brouwers**, for the practical training she gave me during the initial stage of my project. Her professional and personal advice made the two years to pass with ease. I will both miss and never forget her motherly love and affection.

I thank **Dr. Germán Avila-Sakar's** who allowed me to use his dissection microscope, without which I would still be dissecting those fly testes. I would like to thank **Dr. Susan Lingle, Dr. Craig Willis, Dr. Anne Adkins, Dr. Ed Byard** and **Dr. Lisa Sinclair** for their encouragement and support.

I am grateful to all the technicians in the Department of Biology, for assisting me in many different ways to finish this research project in two years. **Dr. Beata Biernacka** deserves a special mention for her effort in improving my presentation skills through my experience in teaching undergraduate students.

Although, I have an exhaustive list of friends to thank individually, due to space constraints I will thank some of my friends who helped me throughout the program. My special thanks to my best friend **Nada Sagga,** for her kind advice, support and constant appreciation. **Naseta Zarin**, **Siavash Darbandi**, **Felix Martinez Nuñez**, **Adriana Suarez Gonzalez, Maram Felemban, Jaimee Dupont** and **Kristin Jonasson** were among those who kept me going from the beginning through their love and support. I appreciate my friend **Scott Finn**, for all the effort he made to improve my outward interests. My acknowledgments stretch to friends back home, always present in my thoughts and heart, for their care and support. To all my friends, regardless of where you live, I greatly value your friendship.

I wish to thank my entire extended family especially my grandfather **Arumugam Subramanium**, my uncle **Thiruvalluvar Muthusamy**, my aunt **Gomathi Thiruvalluvar**, my brothers **Senthil Kumar**, **Ramkumar** and **Sivakumar**, for their unconditional love, encouragement and support throughout this endeavor.

I cannot finish without saying how I feel honored to thank my parents, **Sundararajan Karuppasamy** and **Ramalakshmi Sundararajan** who taught me the value of hard work. They bore me, raised me, supported me and loved me. To them I dedicate this thesis.

Table of contents

List of Tables

List of Figures

List of Abbreviations

DDH_2O	-	Double distilled water
CO_2	-	Carbon dioxide
DNA	-	Deoxyribonucleic acid
RNA	-	Ribonucleic acid
cDNA	-	Complementary deoxyribonucleic acid
°C	-	Degree Celsius
µl	-	Microliter
µM	-	Micromolar
fmol	-	Femtomolar
nm	-	Nanometer
ml	-	Milliliter
cm	-	Centimeter
mm	-	Millimeter
g	-	Grams
mg	-	Milligrams
min	-	Minutes
sec	-	Seconds
F1	-	Filial 1
TBE buffer	-	Tris-Boric acid-EDTA buffer
PCR	-	Polymerize chain reaction
qRT-PCR	-	Quantitative real-time polymerize chain reaction
Ct	-	Threshold cycle

ANOVA	-	Analysis of variance
Taq polymerase	-	Polymerase isolated from *Thermus aquaticus*
mA	-	Milliampere
W	-	Watts
SDS	-	Sodium dodecyl sulfate
Tris	-	Tris (hydroxymethyl)aminomethane
DTT	-	Dithiothreitol
CHAPS	-	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
HCl	-	Hydrochloric acid
APS	-	Ammonium persulfate
TEMED	-	Tetramethylethylenediamine

List of Appendices

1. INTRODUCTION

1.1 Speciation:

Carolus Linnaeus (1707–1778), a Swedish botanist, formulated an orderly system for classifying organisms using a *"binomial nomenclature"* which made a huge impact in the field of taxonomy. He hierarchically classified plants and animals into five major categories: Kingdom, Class, Order, Genus and Species. This system of classification remains a gold standard for classifying all organisms, both extant and extinct. Linnaeus considered species as permanent and immutable entities, while conversely Lamarck and Darwin elucidated that species are constantly evolving. Furthermore, Darwin introduced the concept of phylogenetic branching by which an ancestral species may split into two or more derived species. Although biologists from several streams continue to contribute towards the understanding of the process of evolution, defining "species" and "speciation" remains complicated even today. No single species concept manages to explain speciation in sexually, asexually and dually reproducing organisms.

Amongst several existing species concepts, the Biological Species Concept (BSC) is the most widely accepted definition for describing "species" in sexually reproducing organisms (Coyne and Orr 2004). BSC was proposed by Ernst Mayr in 1942, and it defines species as groups of naturally or potentially interbreeding populations that produce viable and fertile progenies but are isolated from other such groups in terms of reproduction. In other words, species are a population of distinctive fertile individuals that are reproductively isolated from other such populations. The definition of species by BSC is also employed as a legal definition in the Endangered Species Act; United States (Freeman and Herron 2001). Considering reproductive isolation as a suitable criterion for

identifying species distinctively, the BSC provides an exclusive, effective, operational (experimental) and the most agreeable definition by scientists and field biologists, among competing concepts. Reproductive isolation barriers or mechanisms are defined as a group of biological features displayed by organisms which directly or indirectly influence sexual reproduction and help impede gene flow between closely related species. Understanding the mechanisms of reproductive isolation is crucial for evolutionary biologists to address the process of speciation (Coyne 1992). Sexually reproducing organisms display a collection of reproductive isolation barriers that can be broadly classified into two categories: prezygotic (premating) isolation barriers and postzygotic isolation barriers. There also exists a class of intermediary isolating barriers that act after gamete transfer but before fertilization (defined as postmating-prezygotic barriers), which are not discussed in this thesis.

1.2 Pre and postzygotic isolation barriers:

Isolation barriers that prevent the transfer of male gamete (sperm or pollen) into the female reproductive tract of dissimilar species are collectively called prezygotic isolation barriers. In terms of conserving resources on production of wasteful progenies (non-viable, sterile), prezygotic isolation barriers are considered the most economic in that they greatly reduce the probability of gene flow. These barriers are placed in several categories including behavioral, ecological and mechanical forms of isolation. Behavioral isolation mechanisms prevent copulation between dissimilar species either via the display of species-specific courtship patterns or by a lack of cross-attraction of different species. A typical illustration of behavioral isolation is the unique display of flight path, light

pattern and light intensity of male fireflies (Lampyridae) by each species to attract females for mating (Lloyd 1966). Ecological isolation arises as a consequence of related species being separated by difference in their niches (habitat), their breeding season and/or specificity in ecological interactions. In two closely related orchids, *Gymnadenia odoratissima* and *G. conopsea*, differences in floral scent composition and spur length leads to the attraction of different members of Lepidopterans for pollination, with no overlap in pollinator species (Huber et al. 2005). In mechanical isolation, the physical incompatibility of reproductive structures (male and female genitalia) restrains copulation between members of different species. Difference in male genitalic surstyli (musculature) and squeezing movements of the abdomen during copulation reproductively isolates two species of sepsid flies (*Microsepsis eberhardi* and *M. armillata*; Eberhard 2001).

When individuals of the same species mate to produce progenies, the cross is biologically referred to as a conspecific cross, whereas when individuals of two dissimilar species mate, the cross is referred to as a heterospecific or an interspecific cross. Isolation mechanisms that act after the union of gametes from a heterospecific cross are grouped together as postzygotic isolation barriers. Hybrids are generated after heterospecific zygote formation and consequently postzygotic isolation barriers act on the hybrids to reduce their fitness relative to the pure parental species. These postzygotic barriers are further classified into extrinsic and intrinsic factors. Some forms of extrinsic postzygotic isolation reduce the fitness of hybrids for ecological reasons; hybrids benefit from normal development, including gametogenesis, but suffer decreased viability or lowered fertility due to lack of a suitable niche (Coyne and Orr 2004). Hybridization between two closely related butterfly species, *Heliconius melpomene* and *H. cydno* is rare

in nature; hybrids produced are non-mimetic and suffer low mating fitness (Jiggins et al. 2001).

Intrinsic postzygotic isolation barriers lead to the generation of hybrid progenies with inborn developmental defects resulting in inviable or partially/completely sterile offspring. Hybrid inviability occurs when hybrids generated from heterospecific mating fail to survive and die at any developmental stage of the life cycle. Hybrid inviability can be the result of problems associated with unifying two divergent developmental and genetic systems into a single genome. Sturtevant (1920) observed that when *Drosophila melanogaster* females are crossed with *D. simulans* males, only hybrid female progenies survived; hybrid males die at the larval stage. The reciprocal cross only yielded hybrid males while hybrid females die as larvae; all progenies from both crosses were sterile (Sturtevant 1920). Hybrids generated between closely related species of *Rana pipens* complex show various degrees of morphological defects (reduced heads, absence of mouth, abnormal eyes) ultimately contributing to hybrid inviability (Moore 1946).

Hybrid sterility is another form of postzygotic isolation where hybrid progenies from heterospecific mating undergo partial or complete sterility. Hybrids that fail to produce functional gametes (sperm or eggs) are referred to as having physiological hybrid sterility. Hybrid sterility, one of the best explored reproductive isolation mechanisms, was first described in mules by Aristotle in 350 B.C. (Taylor and Short 1973). A large body of work performed in the 20th century by several biologists including Sturtevant (1920), Haldane (1922), Dobzhansky (1934), Coyne (1984) contributed to the understanding of hybrid sterility mainly employing *Drosophila* as a model system.

1.3 Hybrid male sterility – Theories and explanations:

Over 80 years ago J. B. S. Haldane made a striking observation that among the offspring of interspecific crosses, the sex which carries two different types of sex chromosomes (heterogametic sex) is always more severely affected (Haldane 1922). Haldane's statement has since been referred to as "Haldane's rule" and has been found to be obeyed in taxa where males are heterogametic (Dipterans, Mammals) and in taxa where females are heterogametic (Lepidoptera, Aves), with a few exceptions (Coyne 1985; Licht and Bogart 1985; Sawamura 1996). Hybrid sterility appears to be connected to the sex chromosomes and not to sex itself (Haldane 1922; Craft 1938). The rule depicts an initial stage of reproductive isolation where heterogametic hybrids are initially perturbed, with homogametic hybrids affected in crosses between more distantly–related species (Coyne and Orr 1989; Coyne and Orr 2004). Haldane's rule drew the attention of many scientists because it seemed to apply in all organisms having sex chromosomes with a few exceptions (Coyne, et al. 1991).

There are four main widely accepted genetic explanations for Haldane's rule: The dominance theory, the faster-male theory, the faster-X theory and the meiotic drive theory. The dominance theory (also referred as the Dobzhansky-Muller theory or X-autosome imbalance theory) states that heterogametic hybrids are affected by all X-linked genes involved in genic incompatibilities with autosomes (i.e. 1X: 2A) whereas homogametic individuals are only affected by dominant genes. Homogametic hybrids carrying a complete set of autosomes and an X chromosome from each parental species can mask the effect of recessive X-linked genes causing hybrid problems and the X-autosome balance remains unchanged. The faster-X theory predicts that compared to

autosomal genes, X-linked sterility genes tend to evolve at a faster rate if favorable mutations are partially or completely recessive (Charlesworth et al. 1987; Coyne and Orr 2004). Introgression studies in *Drosophila* that aimed to measure the density of X-linked versus autosomal genes causing hybrid problems yielded mixed results and thereby reduced support for the faster-X theory (Coyne and Orr 2004). Meiotic drive is a force acting on individuals of the heterogametic sex in which two kinds of gametes are generated, which differ in the frequency of alleles they carry at the meiotic drive locus. This will distort Mendelian segregation ratios and alter allele frequencies in a population (Sandler L and Novitski E 1957; Sandler et al. 1959). Meiotic drive often occurs in males carrying X chromosomes, (e.g. *segregation distorter*, *sex-ratio*) such that they produce an excess of female offspring due to the abnormal Y chromosome behavior during meiosis II and spermiogenesis (Montchamp-Moreau and Joly 1997; Cazemajor et al. 2000; Jaenike 2001). Cytological studies demonstrate that Y chromosome-bearing spermatids fail to individualize properly. In heterogametic hybrids, the presence of sex chromosomes from different species creates an imbalance in distorter elements which can result in unisexual sterility (Frank 1991; Hurst and Pomiankowski 1991). Although meiotic drive theory might explain heterogametic hybrid sterility, it fails to explain how it can act as an evolutionary force causing postzygotic isolation (Coyne and Orr 2004).

Studies in *Drosophila* showed that reproductive tissues (testis and accessory gland) are more divergent than non-reproductive tissue, suggesting that the male reproductive system including spermatogenesis could be intrinsically disrupted in heterospecific hybrids (Coulthart and Singh 1988a, 1988b; Thomas and Singh 1992). These studies have lead to the creation of a new hypothesis, “faster-male theory”, stating

that hybrid male sterility is accelerated greatly and evolves faster than hybrid female sterility in male-heterogametic taxa (Singh 1990; Wu et al. 1993, 1996). The theory argues that a strong selective pressure could be exclusively observed on male reproductive characters like genital morphology and seminal fluid proteins, making spermatogenesis susceptible to perturbation in sterile male hybrids (Wu and Davis 1993). Investigation of sterility factors through introgression between the species in the *Drosophila simulans* clade reported that a large number of male sterility factors evolve more rapidly than either female sterility or inviability genes (Hollocher and Wu 1996; True et al. 1996; Tao et al. 2003a, 2003b; Tao and Hartl 2003). Recent microarray studies in *Drosophila* confirm misregulated (mainly down regulated) genes were mainly affecting male hybrids' reproduction and they were under expressed relative to the parental species (Michalak and Noor 2003; Moehring et al. 2007). All of the above studies provide support for the faster-male theory as a major contributor to hybrid male sterility although they fail to explain hybrid inviability.

1.4 Phenotype of hybrid male sterility in *Drosophila*:

In *Drosophila*, the majority of heterospecific crosses produce sterile male and fertile female progenies (Bock 1984). Sterile hybrid male progenies are observed to have characteristically unique disturbances (phenotypic and genetic) in the male reproductive system, mainly testes and in the sperm developmental pathway (spermatogenesis) compared to any other developmental processes (Sturtevant 1920; Lancefield 1929; Dobzhansky 1933, 1934; Kulathinal and Singh 1998). Initial studies using different races of *Drosophila sp*. (later identified as two different species, *Drosophila pseudoobscura*

and *D. persimilis*) reported that testes in the F1 hybrid are rudimentary, smaller and showed spermatogenic aberrations compared to the parental testes (Lancefield 1929; Dobzhansky 1934; Dobzhansky 1970). Cytological investigation in the F1 testes revealed that the second meiotic division is absent and abnormal spermatids are produced (Dobzhansky 1934). Further studies using transplantation methods, where testes were transplanted from hybrid larvae into parental species, showed that transplanted adults became sterile. All of the above results suggest that the defects in hybrid testes are autonomous, determined by their own genetic constitution and not by interactions between gonadal and surrounding tissues (Dobzhansky and Beadle 1936).

Interspecific hybridization tests done between four closely related species of the melanogaster complex (*Drosophila melanogaster*, *D. simulans*, *D. mauritiana*, *D. sechellia*) showed variation in the morphology of the testes despite the fact that all hybrid males were sterile (Lachaise et al. 1986). They further classified the sterile testes into 4 different types based upon external morphology and presence and motility of sperm (normal testes with amotile sperm, aspermic normal testes, one atrophied testis but aspermic and both atrophied aspermic testes). Initial cytological analyses to understand spermatogenic defects in the hybrid testes specifically focused on two stages on spermatogenesis: early spermatid stage (onion cell stage) and the sperm bundle stage (Perez et al. 1993). Hybrids generated from two heterospecific crosses (using *D. simulans* as female and *D. mauritiana*, *D. sechellia* as male) produced normal early spermatids but sperm bundles failed to develop normally. Similar spermatogenic phenotypes in hybrid testes were observed using introgression analysis (Perez et al. 1993; Cabot et al. 1994). A further intense cytological study using light and electron microscopy in six possible

interspecies hybrids of the *D. simulans* clade (*D. simulans*, *D. mauritiana* and *D. sechellia*) reported two distinct classes of spermatogenesis arrest phenotypes: premeiotic or postmeiotic varying by the direction of the cross (Kulathinal and Singh 1998). Cross-sectional examination of premeiotically defective testes (e.g. In F1 hybrids generated from *D. mauritiana* (female) × *D. simulans* (male)) showed two to eight mitotic cyst cells, devoid of sixteen cell cysts and sperm bundles. The reciprocal cross showed testes having underdeveloped axonemal complexes, loosely packed spermatids with copious cytoplasm in a preindividualized state (postmeiotic arrest).

1.5 Genetics of hybrid male sterility in *Drosophila*:

Consistent abnormalities affecting fecundity of the hybrid progenies but never observed in the parental species suggest the existence of complex, improper genetic interactions exclusive to hybrids. The genetic basis of hybrid male sterility could be either due to chromosomal effects (chromosomal sterility) or due to genetic effects (genic sterility). Chromosomal rearrangements during meiosis and chromosomal pairing problems that might cause hybrid sterility are not discussed in this chapter. A great deal of genetic studies has identified putative genes causing sterility in hybrid *Drosophila* through introgression analysis monitored by visible mutations or allele specific DNA markers. Backcross analysis from hybrids between *D. simulans* and *D. mauritiana* has mapped the *forked* allele to the X chromosome and the presence of the *forked* allele in a different species chromosomal background results in sterility (Coyne 1984; Coyne and Charlesworth 1986). High resolution genetic mappings on the same region with additional markers led to the identification of a gene, *Odysseus (Ods)*, which causes

sterility (Perez et al. 1993). Demarcation of the *Ods* locus using gene cloning have determined that *Ods* as having two exons including a homeobox motif (i.e. DNA binding) and named its transcript *OdsH* (Ting et al. 1998). Further investigation of this locus revealed that complete sterility was observed only when cointrogressed with adjacent gene segments (Perez and Wu 1995). Investigations to estimate the density of hybrid male sterility factors using QTL mapping and introgressions identified two remarkable findings in hybrids generated between *D. simulans* and *D. mauritiana* (Tao et al. 2003a, 2003b). Firstly, the density of hybrid male sterility factors on the X chromosome is approximately 2.5 times the density of those factors in autosomes. Secondly, QTL analysis identified 19 third chromosome loci with a complex pattern of epistatic interaction capable of causing hybrid male sterility.

Heterospecific crosses between *Drosophila pseudoobscura bogotana* (females) and *D. pseudoobscura pseudoobscura* (males) generate incompletely sterile F1 males (they become weakly fertile when aged) and produce almost all daughters (a sex-ratio distortion) due to meiotic drive (Orr and Irving 2005). A recent study in this hybrid identified that *GA19777*, which was renamed as *Overdrive* (*Ovd*), affects both segregation distortion and hybrid male sterility (Phadnis and Orr 2009). Although the genetics of hybrid male sterility has been greatly explored, very few studies have tried to elucidate what occurs after transcription. Studies on protein divergence using two-dimensional gel electrophoresis on reproductive (testes and ovaries) and non-reproductive tissues (brains, malpighian tubes, wing discs) among species of the melanogaster complex and virilis group revealed a higher rate of protein divergence amongst reproductive tract proteins than proteins involved in non-reproductive functions

(Coulthart and Singh 1988; Thomas and Singh 1992; Civetta and Singh 1995). Two dimensional protein profiles of testes of parental and hybrid flies (*Drosophila simulans* and *D. sechellia*) showed that most parental proteins (97.4%) were present in hybrids. Only 8% of the total genes expressed were unique and possibly involved in hybrid male sterility. These results suggest that only a few genes affecting gametic development were sufficient to produce hybrid male sterility (Zeng and Singh 1993).

All of the above studies established a link between physiological, morphological (cytological abnormalities) and genetic divergence in testes-expressed genes. Thus, it is fair to assume that genes controlling the process of sperm development (spermatogenesis) might be primary targets of disruption in sterile hybrid males.

1.6 Spermatogenesis:

Spermatogenesis is a multistep biological process by which a single germline stem cell undergoes a sequence of divisions and cellular and morphological changes to become mature motile sperm. The cellular differentiation and genetic basis of normal spermatogenesis is well characterized and extensively studied in *Drosophila melanogaster* (Fuller 1993; Fuller 1998; Wakimoto et al. 2004; White-Cooper et al. 2009). Spermatogenesis occurs in a pair of long coiled tubular testes which measures about 0.1mm in diameter and 2.0 mm in length (Lindsley and Tokuyasu 1980). Each testis is closed at the apical end and distally connected to the seminal vesicle, where mature sperm are coiled and stored until transfer to the female.

In *Drosophila*, spermatogenesis can be broadly divided into four sequential stages: 1. Initial germ-line proliferation 2. Four rounds of mitotic amplification 3. Two

rounds of meiosis 4. Final stage of spermatid differentiation (spermiogenesis). The germ-line proliferation center which is located in the apex of the testis comprises three types of cells: apical cells, germ-line stem cells and cyst progenitor cells (Fuller 1993). Groups of densely packed apical cells which form the hub of the center are firmly attached to the apex of the testis. About 6-8 germ-line stem cells are arrayed around the hub cells, each associated with two cyst progenitor cells. Both apical and cyst progenitor cells are somatically derived. Spermatogenesis is initiated by the division of one germ-line stem cell into two daughter cells. The daughter cell that is displaced away from the apical hub matures into 64 motile sperm, whereas the daughter cell that is attached to the hub remains as a stem cell. The germ-line stem cell is accompanied with two non-dividing cyst progenitor cells and forms a cyst, the central core of sperm development.

Each germ-line stem cell undergoes 4 rounds of mitotic proliferation and forms a cyst of 16 immature primary spermatocytes (Fig. 1). Primary spermatocytes are diploid cells interconnected by cytoplasmic bridges due to incomplete cytokinesis. The spermatocytes undergo a brief maturation period of about 90 hrs. At this phase, spermatocytes grow 25-fold in cell volume and undergo morphological changes that distinguish early from mature primary spermatocytes (Fuller 1993; Lindsley and Tokuyasu 1980).

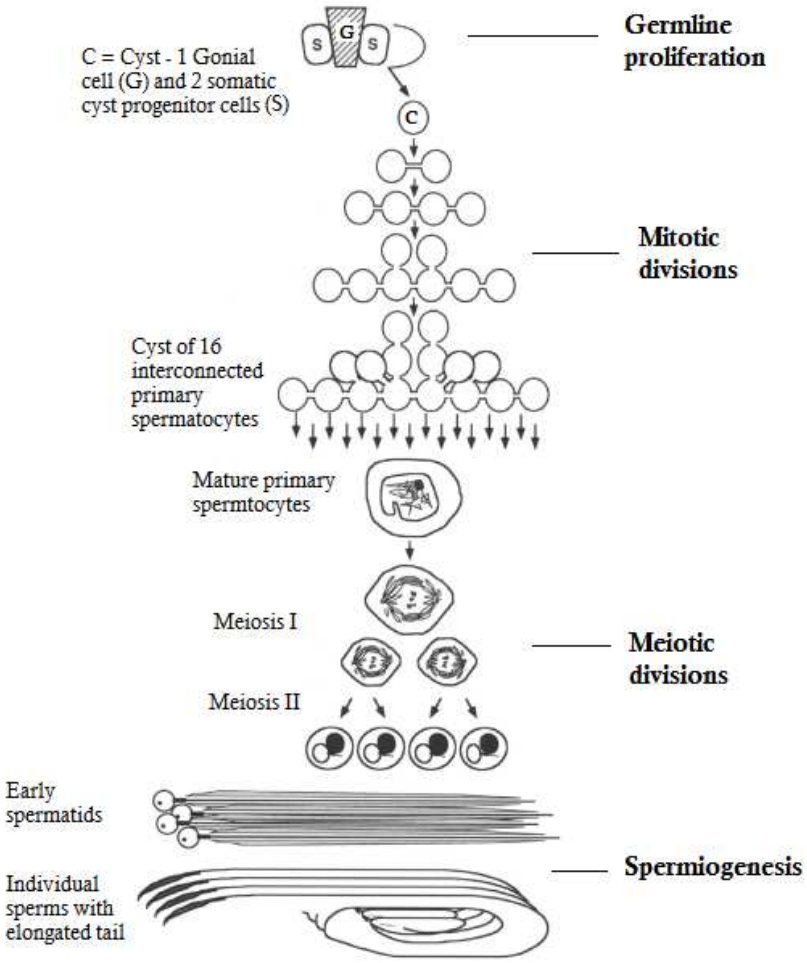

Figure 1.1. Overview of spermatogenesis. (Adapted and modified from Fuller 1998)

Mature spermatocytes are the largest spermatocytes in the testis which are clearly identified by their prominent large, centrally located nuclei viewed by light microscopy (Fuller 1993). Autoradiographic studies have demonstrated that almost all transcription ceases before the mature spermatocytes enters into the next stage (Olivieri and Olivieri 1965; Gould-Somero and Holland 1974). Genes required in meiosis and spermiogenesis are transcribed pre-meiotically and stored in primary spermatocytes; however, a few genes were recently detected to be transcribed post-meiotically (Barreau et al. 2008; White-Cooper 2010). After mitosis, mature spermatocytes undergo 2 rounds of meiosis

which occurs in rapid succession and form 64 haploid spermatids. During the onset of meiosis, the nuclei of the primary spermatocytes become spherical with distinct nuclear envelopes. Mitochondria appear as dark bars and align parallel to the nucleus on the equatorial region of the spindle. Meiosis I is a reduction division in which the diploid chromosomes reduce by half. Shortly after a brief interphase the spermatocytes enter meiosis II which is a simple cell division. The final product of meiosis II is a cyst of 64 interconnected spermatids with two non-dividing cyst-progenitor cells.

Spermiogenesis lasts approximately 134 hours. The interconnected spermatids undergo maturation, elongation and individualization and transform into motile sperm (Lindsley and Tokuyasu 1980). All mitochondria in the spermatids aggregate to form two giant mitochondria. Both mitochondria are interleaved and forming tightly wrapped layers adjacent to the nucleus. Mitochondria at this stage are referred to as Nebenkern and resemble an onion in cross section. The axoneme elongates and consists of microtubules that comprise the basic 9+2 architecture (nine outer doublet microtubules surrounding a central pair of singlet microtubules; Fuller 1993). The spermatid nucleus undergoes a complete morphological change from a spherical to a thin needle shaped structure. By the end of the elongation phase, the very condensed, needle-shaped nucleus measures about 9 μm in length and about 0.3 μm in width, including the acrosome. During the individualization process, enlargement of the spermatid bundle called the cystic bulge progresses along the entire length of the bundle from head to tail. Most of the spermatid cytoplasm and intercellular bridges are lost by the cystic bulge. Excessive nuclear envelope, cytoplasmic organelles and minor mitochondrial derivatives are expelled caudally from the cystic bulge. At the end of individualization, each resulting

spermatozoa has the acrosome which is firmly attached to the plasma membrane and the nuclei attached to the acrosome. Both the cyst cells differentiate, distinguishing one from each other. The head cyst cell forms a cap over the spermatid nuclei while a tail cyst cell surrounds the remainder of the cyst and encapsulates the elongating tail. After individualization, the sperm bundle coils, starting from the head and changes its linear confirmation into a coiled confirmation. Abnormal sperm that fail to individualize are pushed caudally and accumulate into a waste bag at the caudal tip of the bundle. Completely developed sperm are liberated from the cyst cells into the testicular lumen while abnormal sperm in the waste bag are subjected to lysosomal degradation. Mature sperm tend to align in the seminal vesicle and are passed to the female during copulation. Approximately 28 hours after hatching, a larval testis has completed primary spermatogenesis and sperm bundles are found in completely grown larvae. The testis of the early pupa (24-30 hours after the onset of pupation) has all the successive stages of spermatogenesis (Cooper 1965).

1.7 Gene expression in sterile hybrid males:

Advances in molecular tools and techniques have greatly facilitated research in genetics and evolution to better understand the process of speciation and reproductive isolation mechanisms. Recently, genomic and proteomic tools have been highly employed to identify gene interactions or misregulation contributing to hybrid sterility in *Drosophila*. Genome wide expression profiling (microarray) of *Drosophila simulans*, *D. mauritiana* and F1 hybrid (*D. simulans* (female) × *D. mauritiana* (male)) using microarray assays have reported that genes involved in spermatogenesis were mainly

down regulated in the hybrid (Michalak and Noor 2003, 2004; Ranz et al. 2004; Michalak and Ma 2008). Similar microarray studies using three species in the *D. simulans* clade (*D. simulans*, *D. mauritiana* and *D. sechellia*) have also reported misexpression of spermatogenesis genes and other male specific genes in the sterile hybrids (Haerty and Singh 2006; Moehring et al. 2007). In spite of their genome-wide analysis, these microarray results are to be regarded with caution due to two main reasons. Firstly, hybridization bias due to sequence divergence between *D. melanogaster* genomic arrays (cDNA probes) to hybridize samples isolated from hybrids sired using other species (*D. simulans*, *D. mauritiana*, and *D. sechellia*) could have resulted in spurious data. Secondly, no quantitative difference (significant fold increase or decrease) of expression was noticed between pure and hybrid species (Michalak and Noor 2003). A better approach has been the use of custom sperm array analysis (microarray platform of genes involved in spermatogenesis) with species-specific alleles (Moehring et al. 2007). However, this study used RNA extracted from whole flies and so it decreased the ability to detect testes-specific gene transcript differences. These reasons exemplify the need to employ better techniques, such as real-time PCR, that do not randomly scan entire genome sequences and target expression pattern analysis of the tissues affected in the male sterile hybrid (testes). Moreover, no studies have identified the expression pattern of genes in hybrid fertile females that are commonly expressed in male and female gametogenesis.

The detection of quantitative differences in gene expression can be better detected through fine resolution and highly sensitive techniques such as quantitative real-time PCR which can quantify the amount of messenger RNA being synthesized in a cell

irrespective of tissue type and time. A recent study used real-time PCR to identify the gene expression pattern in *D. simulans*, *D. mauritiana* and F1 hybrid (*D. simulans* (female) × *D. mauritiana* (male)) using RNA isolated from testes and whole body (Catron and Noor 2008). Four candidate genes, two expressed during meiosis (*aly* and *comr*) and two spermiogenesis or post-meiotic (*don juan* and *Mst84D*) genes were assayed. The authors found that post-meiotic genes were down regulated in sterile hybrid males relative to parents and the under expression was significantly lower in whole bodies' samples than testes. However, this study failed to sample the expression pattern of genes prior to meiosis (germline proliferation and mitosis).

1.8 Objectives:

Following are the objectives of my study to address the hypothesis "Is sterility a consequence of post-meiotic gene misregulation during spermatogenesis in hybrid *Drosophila* males?"

1. To determine whether the expression of candidate genes, selected from all stages of spermatogenesis, are down regulated in sterile interspecies hybrids' testes compared to testes from parental species.
2. To test whether the expression of down regulated genes is testes-specific.
3. To verify that the down regulation of candidate gene expression is restricted to interspecies sterile hybrids rather than hybrids in general.
4. To verify that the down regulation of candidate gene expression is unique to the male pathway of gametogenesis and not shared with the fertile hybrid females' gametogenesis (oogenesis).

2. MATERIALS AND METHODS

2.1 *Drosophila* stocks:

Three species of *Drosophila* were used for research during the completion of this thesis. Wild type *Drosophila simulans* California strain (sim2) and Congo strain (isofemale 15) were kindly provided by Dr. Andrew G. Clark (Cornell University). *D. mauritiana* strain (14021-0241.01) was obtained from the *Drosophila* species stock center (UCSD, La Jolla, CA). Stocks were reared in cylindrical polypropylene bottles (237 ml) and cylindrical vials (28.5mm × 95mm) containing fresh cornmeal-molasses-yeast-agar (CMYA) medium (Appendix I). All flies were maintained in an incubator at 20º C with 12 hour light-dark cycle. Adult flies were placed in bottles containing fresh media, and after twelve to fourteen days adult flies were dumped. Newly emerging adult flies were transferred into new bottles. Stocks were maintained in this fashion throughout the research to assure healthy stocks and to prevent overlapping of generations.

Interspecies hybrids were generated by crossing females from *D. simulans* (California strain) with *D. mauritiana* males. Repeated trials to generate hybrids from the reciprocal cross did not yield any progeny, in agreement with previous finding of strong cryptic barriers to hybridization between *D. mauritiana* females and *D. simulans* males (Price et al. 2001). Intraspecies hybrids were generated by mating females from *D. simulans* (California strain) with *D. simulans* (Congo strain) males. In order to generate these hybrids, stock bottles were emptied late in the afternoon making sure no adult flies remained in the bottles. Virgin female flies were collected the next morning because the majority of flies emerge from pupae at dawn. Newly emerged flies are easily

distinguished by their pale pigmentation and any flies with dark pigmentation were discarded as they might not be virgins (Greenspan 1997). Virgin flies were collected every morning, and flies eclosed after each collection were dumped every night for next day morning collection. Virgin flies were lightly anesthetized with CO_2 and virgin males were separated from virgin females. Virgin flies were maintained in polypropylene vials containing fresh CMYA media, with no more than 50 flies per vial and aged for three to four days to sexual maturity (Greenspan 1997). Ten virgin males were crossed with 10 virgin females to generate a hybrid generation. After twelve to fourteen days, parental flies were dumped and dental rolls were inserted into media to maximize dry substrate for pupation. F1 hybrid male flies were collected, aged 4-6 days and used for gene expression and proteomic studies.

2.2 Candidate gene selection:

One thousand and eight genes in the FlyBase database (http://flybase.org/) were found to be annotated in *D. melanogaster* under the Gene Ontology term “reproduction”. A survey of these candidate genes narrowed the number down to 145 genes - 96 genes under the term “spermatogenesis” and 49 genes under the term “spermatid development”. All genes under these two terms were individually examined for DNA sequence availability, protein function and literature support, because many genes annotated were not well characterized. Fifty two genes from both ontologies were grouped into one of the four major developmental stages in the spermatogenesis pathway: Germline proliferation (4 genes), mitosis/mitotic arrest (8 genes), meiosis/meiotic arrest (26 genes) and

spermatid development/spermiogenesis (14 genes). Two candidate genes from each stage were randomly selected as representing each stage of sperm development (Table 2.1).

2.2.1 Candidate gene selection: Detailed protocol:

1. Open FlyBase home page and click “TermLink”

2. Under “hierarchy structures” categories select “Biological Process” and click on “reproduction” to view the spanning tree.

3. Scroll down and click “Genes” in the “records annotated with this term or any of its children terms” to display all the genes annotated under the term reproduction

4. Click “Result analysis/Refinement” and select “Biological process” in the pop-up box to list all genes grouped under specific biological terms.

5. Search for “spermatogenesis; GO: 0007283” and “spermatid development; GO: 0007286” and click on the respective “Related records” link to view genes classified under these ontologies.

6. Click on each gene to retrieve complete information and its sequence.

Table 2.1 Candidate genes selected for gene expression analysis

Stages in spermatogenesis	Candidate gene	Gene symbol	Function	References
Germline proliferation	hopscotch	*hop*	A JAK tyrosine kinase involved in the JAK-STAT signaling pathway	Kiger et al. (2001); Silver and Montell (2001); McGregor et al. (2002)
	piwi	*piwi*	Germline stem cell renewal of both the sexes	Cox et al. (2000); Wong et al. (2005)
Mitosis	bag of marbles	*bam*	To cease proliferation of amplifying mitotic cells or promote their entry into meiotic cell cycle	McKearin and Spradling (1990), Gonczy et al. (1997)
	benign gonial cell neoplasm	*bgcn*		
Meiosis	cannonball	*can*	Progression through the meiotic cell cycle and the onset of the spermatid differentiation program	Lin et al. (1996); Hiller et al. (2004)
	spermatocyte arrest	*sa*		
Spermiogenesis	fuzzy onions	*fzo*	Encodes for transmembrane GTPase required for mitochondrial fusion during Nebenkern formation	Hales and Fuller (1997); Hwa et al. (2002); Mozdy and Shaw (2003)
	Bruce	*Bruce*	Ubiquitin conjugating enzyme acts to protecting excessive caspase activation and death in spermatid nuclei	Cagan (2003); Arama et al. (2003)

2.3 DNA primer design:

Two reference genes (house keeping genes), *RpL32* (Ribosomal protein L32) and *Act5C* (Actin 5C), were used in this thesis to measure relative expression of candidate genes. The entire coding region of *Drosophila melanogaster* for all candidate genes and two reference genes were obtained from the FlyBase database and saved in FASTA format. These sequences were aligned using the sequence similarity search algorithm BLAST (Basic Local Alignment Search Tool), against *D. simulans* and *D. sechellia* genome (http://flybase.org/blast/). Orthologous hits of *D. simulans* and *D. sechellia* gene sequences were retrieved from the FlyBase database. *D. melanogaster*, *D. simulans* and *D. sechellia* sequences were re-aligned using the local alignment algorithm ClustalW2 (Larkin et al. 2007). Sequence regions with minimal nucleotide differences in the alignment were selected to be targeted for primer design.

All primers were designed using an online primer design program called BatchPrimer3 (You et al. 2008). All primers were designed to amplify products with: melting temperatures greater than 60ºC, GC contents of 50-60% and amplicons of 120-250 bps. Putative secondary structures formed by designed primers were minimized using Oligo Calc, an online secondary structure prediction tool (Kibbe 2007). Primers for *bam* and *RpL32* were designed to span an intron such that any DNA contamination in RNA extraction would result in a bigger product than expected. Information about the designed primers is summarized in Table 2.2 and these primers were purchased from Invitrogen.

Table 2.2. Primer description and sequences.

Primer Description	**Sequence (5’ to 3’)**
*RpL32*_Forward_primer	TACAGGCCCAAGATCGTGAA
*RpL32*_Reverse_primer	ACCGTTGGGGTTGGTGAG
*Act5C*_Forward_primer	CCGTGAGAAGATGACCCAGA
*Act5C*_Reverse_primer	CGGTCAGGATCTTCATCAGG
*hop*_Forward_primer	TGAGTGTGGAGCGTTTGAAG
*hop*_Reverse_primer	TGGACAGAGTGTTGGTGGAA
*piwi*_Forward_primer	ATTGCGAAGAGCACACGAG
*piwi*_Reverse_primer	CCCGTCCCGATAAAATACG
*bam*_Forward_primer	CGCAATCGAAACGGAAAC
*bam*_Reverse_primer	CGGCACCAGACAAAAGGA
*bgcn*_Forward_primer	ACGGTGGCAATAACGGAAC
*bgcn*_Reverse_primer	CGGAATGTGCAAGGGAAC
*can*_Forward_primer	TTCGCTTGTGGTGCCTTC
*can*_Reverse_primer	CTCTCGCCGTACAATCATCC
*sa*_Forward_primer	AAAGCACCGGAGACACAAGA
*sa*_Reverse_primer	CCTGGAAATGGTGGCAAA
*fzo*_Forward_primer	AAGCTCTCGCGTCCAAATC
*fzo*_Reverse_primer	CGCCGTGGAATAAACACCT
*Bruce*_Forward_primer	TTGCCGGAACTTGGATAGG
*Bruce*_Reverse_primer	TCCGCTGCCTGTGTTAATG

2.4 Tissue dissections:

Lysis solution was prepared by adding 5 µl of 2-mercaptoethanol to 500 µl Buffer RLT plus (with RNeasy plus mini kit from Qiagen) and chilled on ice prior to dissection. Needles, forceps, insect pins and microscopic slides to be used for dissection were wiped with ELIMINase solution (Decon Labs, Inc) using Kimwipes to eliminate any DNase, RNase and DNA contamination. About 1 ml of 1X PBS (Appendix II) was placed on a clean microscopic slide. Testes were dissected from 4-6 days old *D. simulans* (California), *D. simulans* (Congo), *D. mauritiana* and the two hybrids. A single male fly was placed on its lateral side in a drop of 1X PBS. The thorax of the fly was poked with a dissecting needle and the abdomen was gently torn open with a Dumont #5 forceps (Fine Science Tools) to expose the contents of the abdomen to saline. A pair of testes with seminal vesicles, identified as a bright yellow-spirally coiled organ was separated from the reproductive tract and other body tissues. Testes with seminal vesicles were then carefully lifted using 0.25 mm diameter insect pin (Fine Science Tools) and gently dropped into the lysis solution. Accidentally punctured testes could be easily identified by a white cloud of sperm that quickly spreads through the saline buffer. Care was taken not to puncture the testes wall. Ovaries were dissected from 4-6 day old parental and hybrid female flies using a technique similar to that described for testes dissection. Each testes and ovary sample consisted of 25 tissues, placed in a 1.5 ml eppendorf tube containing 500 µl of lysis solution. The tubes were stored at -80ºC until RNA extraction. I also stored whole bodies of 25 parental and hybrid males in similar fashion for RNA extraction.

2.5 RNA extraction and cDNA synthesis:

RNA was extracted from the dissected tissue samples (testes and ovaries) as well as whole flies with RNeasy plus mini kit (Qiagen) following the manufacturer protocol. Total RNA was finally eluted in 40 µl of nuclease free water and quantified using a Nanophotometer (Implen, Inc). The concentration of RNA was determined by the absorbance at 260 nm using the following equation (Nanophotometer user manual v2.0):

$$Conc_{nuc} = A_{260} * Factor_{nuc} * Lid\ factor$$

where $Conc_{nuc}$ is the nucleic acid concentration (ng/µl), A_{260} is the absorbance of nucleic acids, $Factor_{nuc}$ is the substance specific factor for nuclei acids (ng*cm/µl) (dsDNA 50, ssDNA 37, RNA 40) and Lid factor is the dilution factor (5, 10, 50 or 100 times, depending on the used LabelGuard lid).

First strand cDNA synthesis was performed in a MJ PTC-200 Peltier Thermal Cycler using an iScript select cDNA synthesis kit. cDNA synthesis for parental and F1 hybrid RNA extractions were carried out in 200 µl PCR tubes with 4 µl of 5X iScript select reaction mix, 2 µl of Oligo $(dT)_{20}$ primer, 1 µl of iScript reverse transcriptase, variable volume of RNA (concentration of RNA was adjusted for uniform concentration across all samples) and brought to a final volume of 20 µl with nuclease-free water. The reaction mix was incubated at 42°C for 80 min for reverse-transcription and 5 min at 85°C for inactivation of the reverse transcriptase. Reverse transcribed cDNA products were quantified following the protocol as described earlier in this section and stored at -20º C until further use.

2.6 PCR and sequencing reaction:

Candidate gene regions amplified by custom designed primers were sequenced in *D. mauritiana* to confirm that the right product had been amplified during qRT-PCR and to identify whether any nucleotide substitution had occurred at the primer sites. Direct sequencing of PCR products was performed by cleaning and sequencing the amplified products. cDNA extracted from parental testes were used as templates for amplification. PCR reactions were carried out in a MJ research PTC-200 Peltier Thermal Cycler. All PCR reactions were performed in 0.2 ml clear PCR tubes with 1 µl of 10 µM primers (each of forward and reverse), 2.5 µl of 10X buffer, 0.6 µl of 10mM dNTPs, 0.2 µl of Taq polymerase, 2 µl of cDNA sample and brought up to final volume of 15 µl with nuclease-free water. Thermal conditions used for amplification are summarized in Table 2.3.

Amplified PCR products were cleaned using E.Z.N.A. Cycle-Pure Kit (Omega Bio-tek) using the manufacturer protocol to remove inorganic impurities, primer dimers and excess reagents. Clean PCR products were quantified using a Nanophotometer to make sure a template concentration of 25 to 100 fmol of purified PCR product was available for sequencing. Depending on the length of PCR product variable amount of template were used (see "Template preparation" section of the CEQ DTCS Quick Start Kit, Beckman Coulter).

A 20 µl sequencing reaction mix was prepared by adding right aliquot of quantified DNA, 2 µl DTCS Quick start master mix, 1.5 µl 10X sequencing buffer,0.3 µl pellet paint, 1 µl of forward or reverse primer and brought up to final volume with sterile DDH_2O. The reaction was carried out in a thermocycler and thermal conditions are

summarized in Table 2.4. Freshly prepared stop solution (0.4 µl of 0.5M EDTA, 2 µl of 3M sodium acetate, 1 µl of glycogen and 1.6 µl of sterile DDH_2O) was immediately added after amplification to each product. 60 µl of 95% ethanol was added to the reaction mix and centrifuged at 4°C for 15 min at 14,000 rpm. The supernatant was discarded, the pellet was washed twice with 100 µl of 70% ethanol and centrifuged at 4°C for 5 min at 14,000 rpm. The final pellet was air dried for 30 min and resuspended with 40 µl of sample loading solution (GenomeLab DTCS Quick start kit). Samples were then transferred into a sample loading plate and overlaid with one drop of mineral oil. The separation buffer (Beckman Coulter) was added to the wells of the buffer plate. The sample loading plate and the buffer plate were placed into a CEQ 2000XL sequencer and ran overnight. Raw sequence data were retrieved using the CEQ system analysis software and exported into text format. Five gene PCR products (*Act5C*, *RpL32*, *can*, *sa* and *Bruce*) were sent for sequencing to the sequencing facility at the Centre for Applied Genomics (SickKids, University of Toronto). The genes partial sequences can be found in GenBank under accession numbers HQ338082-HQ338091.

Table 2.3. Thermal conditions used for PCR.

Steps	Temperature (° C)	Time
1. Initial denaturation/enzyme activation	95	5 min
2. Denaturation	95	45 sec
3. Annealing	60	30 sec
4. Extension	72	45 sec
5. Repeat step 2 to 4 for 34 more cycles		
6. Final extension	72	15 min

Table 2.4. Thermal conditions used for sequencing reaction.

Steps	Temperature (° C)	Time
1. Denaturation	95	45 sec
2. Annealing	50	45 sec
3. Extension	60	4 min
4. Repeat step 1 to 4 for 34 more cycles		
5. Final extension	60	8 min

2.7 Quantitative RT-PCR:

PCR reactions were carried out in a MJ research PTC-200 Peltier Thermal Cycler. The cDNA extracted from *D. simulans* testes and ovaries were used as templates for amplification. All PCR reactions were performed in 0.2 ml clear PCR tubes with 1 µl of 10 µM primers (each of forward and reverse), 2.5 µl of 10X buffer, 0.6 µl of 10mM dNTPs, 0.2 µl of Taq polymerase, 2 µl of cDNA sample and brought up to final volume of 25 µl with nuclease-free water. Thermal conditions used for amplification are summarized in Table 2.3.

All quantitative RT-PCR reactions were performed using iQ SYBR Green Supermix kit from BioRad and carried out in a MiniOpticon PCR System (Bio-Rad Laboratories, Inc). All reactions were performed in 200 µl low profile BioRad white PCR tubes sealed with clear flat caps. RT-PCR amplifications were performed for all candidate genes with 7.5 µl of iQ SYBR Green Supermix, 0.5 µl of each primer (forward and reverse), 0.7µg of cDNA template and brought up to final volume of 15 µl with nuclease-free water. Primer concentrations for all candidate genes and reference genes were 10 µM and 5 µM respectively; concentrations of primers were optimized to minimize primer dimerization during amplification. To add statistical robustness to the analysis, each testes sample in comparisons between sterile hybrids and fertile parental species was replicated six times (Pavlidis et al. 2003; Allison et al. 2006). Thermal qRT-PCR conditions are similar to PCR thermal conditions as described in Table 2.3, but with two changes.

1. A plate read was inserted after extension step (step 4)

2. Final extension step (step 6) was replaced with a melt-curve analysis: Samples were gradually heated from 55°C to 99°C with a plate read at every 0.5°C intervals, holding for 1 second.

The presence of a single amplification of all candidate genes was confirmed by agarose gel electophoresis using 1X TBE buffer (Appendix III). Samples were loaded to freshly cast 1.5% agarose gel containing 0.5µg/ml of ethidium bromide and run at 120 volts, for 35-45 min. A standard curve method was used to determine the binding efficiency of each primer. qRT-PCR was performed for all candidate genes with undiluted and 10^{-1}, 10^{-2} and 10^{-3} dilutions of cDNA template. Three biological replicates were run and the average Ct values were plotted against the log of the dilution. Except for *bgcn* (92%) all candidate gene primers and both reference gene primers resulted in 100% efficiency.

2.8 Quantitative RT-PCR data analysis:

Expression of all candidate genes was quantified relative to the expression of two reference genes, *RpL32* and *Act5C*. Threshold parameters were set for each candidate genes and for both reference genes by selecting the cycle at which the reaction begins to enter the exponential phase of amplification. Ct values (threshold cycle) were obtained for all genes and imported into a spreadsheet. Relative expression of each gene is quantified using the following formula:

$$\text{Relative expression} = E^{\text{Ct }(ref) - \text{Ct }(gene)}$$

where E is the amplification efficiency of the gene; Ct (*ref*) is the Ct value of reference gene and Ct (*gene*) is the Ct value of candidate gene

Since the amplification efficiency of all genes including the reference genes were close to 2, the above equation was modified to:

$$\text{Relative expression} = 2^{\text{Ct}\,(ref) - \text{Ct}\,(gene)}$$

or

$$\text{Relative expression} = 2^{\Delta \text{Ct}}$$

Statistical analysis was performed using ANOVA using species as a fixed factor. The data were first subjected to an angular transformation to fit the ANOVA assumptions of normality and homoscedasticity or using Mann-Whitney nonparametric pair comparisons (SPSS v12.0). Test results were corrected for multiple test comparisons by using false discovery rate corrections to statistical thresholds. I also compared average differences in gene expression between parental species and hybrids by using an approximate randomization analysis with 25,000 permutations (Manly 1991).

2.9 Protein isolation and quantification:

20 pairs of testes from *D. simulans*, 40 pairs from *D. mauritiana* and 50 pairs from F1 hybrids were dissected from 4-6 day old flies as described in section 2.4 and stored in 100 µl of PS buffer (Appendix IV) at -70ºC until use. Prior to SDS electrophoresis, the tubes containing tissue samples were subjected to repeated flash freeze/thaw (5 times) with liquid nitrogen and a hot water bath (60ºC) to disrupt cells. Samples were centrifuged for 20 min at 14,000 g. The resulting supernatant was transferred to another tube and centrifuged for 10 min at the same speed. The final supernatant was stored at -70ºC until further use. The total concentration of each protein sample was determined through calorimetric Bradford protein method (Bradford 1976)

using Coomassie Brilliant Blue G-250 Dye reagent (1X) and Bovine Serum Albumin as the standard (BioRad Quick start Kit). 5 µl of standards (2, 1.5, 1, 0.75, 0.5, 0.25, 0.125 mg/ml) and protein samples were individually added to 0.6 ml centrifuge tube containing 250 µl of Coomassie dye reagent (1X). Each solution was thoroughly mixed and incubated at room temperature for 30 min. A blank solution was prepared by adding 5 µl of PS buffer to 250 µl of Coomassie dye reagent and used to zero the Nanophotometer before measuring the absorbance of standards and samples. 200 µl of each solution was individually transferred to a 15 mm center beam high, disposable ultra-micro cuvettes (BrandTech Scientific, Inc). To construct a calibration curve, cuvettes containing standards were placed sequentially (starting from the least concentrated standard) on the Nanophotometer (Implen, Inc) and absorbance of each standard was individually measured at a wavelength of 595 nm. The instrument generates a calibration graph showing linear regression between the absorbance of standards to their concentrations. Using the calibration graph, the absorbance of the protein sample was measured. The concentration of the standards plotted in the calibration graph was used as a reference to determine the concentration of the unknown protein samples.

2.10 Preliminary separation of protein using SDS gel electrophoresis:

Casting of the SDS gel and subsequent protein separation through electrophoresis was performed using a Miniprotean Tetra Cell Unit (Bio-Rad Laboratories, Inc). Gel casting frames, glass plates, combs and gaskets were initially sterilized with 95% ethanol and then with DDH_2O. Glass plates were assembled on the casting frame in a “sandwich” following the instructions in the product supplied manual. To distinguish the level until

the separating gel must be poured, a spacer comb was inserted in the glass sandwich and a mark was made about 1 cm below the spacer comb with a permanent marker. Once after the mark was made the spacer comb was removed. Freshly prepared separating gel (Appendix V) was gently poured into the pre assembled glass sandwich, until the marked level avoiding air bubbles. After 2-5 minutes, approximately 1 ml of DDH_2O was overlaid on the gel to prevent dehydration and the gel was allowed to solidify for 2 hours at room temperature. After solidification, water overlay was removed and freshly prepared stacking gel (Appendix VI) was poured until the top (approximately 2.5 ml). A spacer comb was inserted immediately without trapping any air bubble and the gel was allowed to polymerize for 1.5 hours at room temperature.

After complete polymerization, the glass sandwich was assembled in the electrode frame and gently transferred into an electrophoresis tank and 750 ml of 1X tank buffer (Appendix VII) was poured. 20 µg of each protein sample (parental and hybrid) was individually mixed with 10 µl of 2X sample buffer (Appendix VIII) and boiled to 95ºC for 5 min. These samples were loaded into wells and subjected to electrophoresis at 25mA, 250 volts, 10 W for about 2 hours or until the dye front reached the bottom of the gel. Once the run was completed, the gel was removed from the electrophoresis unit and stained using Coomassie blue as described in Appendix IX.

2.11 2D gel electrophoresis:

2.11.1. Protein clean-up:

SDS gel electrophoresis as described in section 2.10 was performed to ensure that all extractions had enough protein to be reliably used for 2D gel electrophoresis. All protein samples were cleaned using the 2-D clean-up kit (GE healthcare) to remove

interfering substances like detergents, salts, lipids, phenolics, nucleic acids, thereby improving the quality of the final gel. About 100 µg of each protein sample was cleaned with the 2-D clean-up kit (GE healthcare) using the manufacturer protocol. Proteins were precipitated during the final step and the pellets were resuspended in 140 µl of the DeStreak rehydration solution (GE healthcare). The volume of the DeStreak rehydration solution varies according to the length of the IPG strips, as 7 cm IPG strips were used for this study, the protein pellets were resuspended in 140 µl of rehydration solution. Resuspended clean protein samples were either stored at -70°C or immediately used for downstream analysis.

2.11.2. IPG strip rehydration:

The 7-cm ready-made IPG strips, Immobiline DryStrip gels (GE healthcare) with a pH range of 4-7 were used for the first dimension. About 125 µl of cleaned parental and hybrid protein samples were evenly pipetted into individual lanes of the rehydration tray (Amersham Biosciences) covering the length of the IPG strip. With the help of a fine pair of forceps, a plastic layer covering the IPG strip was carefully peeled off starting from the anodal end. The IPG strip was laid onto a rehydration tray, with the gel side facing the sample. Care was taken to evenly distribute the sample under the strip and not to trap any air bubbles under the strip. The strip was overlaid with 5-10 ml of DryStrip cover fluid (Amersham Biosciences) and incubated for 10 to 20 hours at room temperature.

2.11.3. First dimension run:

The first and second dimension were run on a horizontal Multiphor II electrophoresis system (Amersham Biosciences). Approximately 10 ml of DryStrip cover fluid (Amersham Biosciences) was evenly pipetted on top of the cooling plate part of the electrophoresis unit. Immobiline DryStrip tray and strip aligner (Amersham Biosciences) were placed on the cooling plate following the instructions in the manufacturer protocol (see section 6.2 of the "Immobiline DryStrip kit for 2-D electrophoresis with immobiline DryStrip and ExcelGel SDS" manual, Catalogue number - 18-1038-63). Rehydrated IPG strips were removed from the rehydration tray and excess cover fluid was removed by touching the edge of the strips with Kimwipes. These strips were carefully placed on the center of the IPG strip aligner with the gel side facing up as well as the positive sign facing the anode end of the electrophoresis system. Two electrode strips (11cm × 1.5 cm) were moistened with DDH_2O and placed at both ends of the rehydrated strips, touching the gel surface. Electrodes were tightly secured over the electrode strips and 70-80 ml of Dry Strip cover fluid was poured to the strip aligner covering the IPG strips and buffer strips. The gels were run at 20°C as indicated in Table 2.5 using EPS 3501 XL power pack (Amersham Biosciences). Once the run was complete, the IPG strips were carefully removed from the rehydration tray and excess cover fluid was removed by touching the edges of the strips not directly touching the gel. These strips were immediately used for second dimension electrophoresis or wrapped with aluminum foil and stored at -80°C until further use.

Table 2.5. Running conditions for first dimension electrophoresis.

Phase	Voltage	mA	W	Time (hrs)	Vh
1	300	1	5	0.01	1
2	300	1	5	4.5	1350
3	2000	1	5	5	5750
4	2000	1	5	6.5	13000

2.11.4. Equilibration of IPG strips:

Immediately prior to second dimension electrophoresis, all IPG strips were individually placed in 10 ml of freshly prepared equilibration buffer no.1 (Appendix X-A) and agitated in a platform shaker for 10 min. Equilibration buffer no. 1 was discarded and 10 ml of equilibration buffer no. 2 (Appendix X-B) was immediately added. The strips were again agitated for 10 min. After the second equilibration, the strips were placed on a piece of filter paper moistened with DDH_2O to drain the equilibration buffer. The strips were left in this position for up to 10 min until the second dimension was assembled.

2.11.5. Second dimension run and staining:

Approximately 15 ml of Dry Strip cover fluid was evenly pipetted on the cooling plate of the Multiphor system and set to 15°C. The ExcelGel SDS 12-14 (GE healthcare) was placed on the cooling plate, with the plastic support facing the plate and care was taken no to trap any air bubbles. A thin plastic film covering the surface of the gel was

removed. Two electrode strips (25 cm × 5 cm; GE healthcare) were used to create buffer strips for the anodic and the cathodic ends; one of the buffer strips was dampened with 10-20 ml of anode buffer (Appendix XI-A) and the other strip were dampened with 10-20 ml of cathode buffer (Appendix XI-B). The cathode and anode buffer strips were placed on the cathodal and anodal edge of the cooling plate over the gel. The equilibrated IEF strips were carefully placed on the ExcelGel following the instructions in the manufacturer protocol (see section 6.3 of the "Immobiline DryStrip kit for 2-D electrophoresis with immobiline DryStrip and ExcelGel SDS" manual). Application pieces (GE healthcare) were used to create space between IPG strips as well as to load the protein marker. Once the electrodes were placed and the assembly was complete, the samples were run as indicated in Table 2.6.

Table 2.6. Running conditions for second dimension electrophoresis.

Phase	**Voltage**	**mA**	**W**	**Time (mins)**
1	1000	20	40	45[a]
2	1000	40	40	5[b]
3	1000	40	40	160[c]

a – When the bromophenol front has moved 4-6 mm from the IPG strips, the IPG strips and application pieces were removed.

b – When the front moved a further 2 mm, the cathode buffer strip and the cathode electrode was moved forward to cover the area of the removed IPG strips.

c – When the bromophenol front has moved into the anodic buffer strip, the run was stopped.

Once the run was complete, the ExcelGel SDS was removed from the cooling plate and stained using PlusOne Silver staining kit using the manufacturer protocol (GE healthcare). The silver stained SDS gel was documented in the gel documentation unit (Bio-Rad Laboratories, Inc) with Quantity One software (Bio-Rad Laboratories, Inc) and the image was saved in (.1sc) format.

2.11.6: Detection of protein spots:

Parental and the hybrid second dimensional gel images (in .1sc format) were individually uploaded to the ImageMaster 2D Platinum software (GE healthcare). After adjusting the contrast of the image, protein spots were detected in all gels following the instructions in the manufacturer manual. All detected spots were then visually inspected to remove false reports. The number of spots in parental and hybrid samples were counted and scored for analysis. About one or two spots clearly distinguishable on all images were correspondingly designated as "landmarks" (following the instructions in the manufacturer protocol), in order to facilitate the matching process. Once landmarks were defined, all images were matched to identify the common spots shared between all the images. The matches are displayed with the match vectors, which links the spots in an image with the corresponding spots in other gel images

To identify a particular protein of interest, protein sequence of *D. melanogaster* of the particular gene retrieved from the FlyBase database (http://flybase.org/) was fed into online Protparam tool (Gasteiger et al. 2005) of Expasy proteomic server

(http://www.expasy.org/). The tool generates a theoretical pI and molecular weight of the protein sequence which was saved for future analysis. Theoretical pI and molecular weight predicted by the Protparam tool is similar to the annotated polypeptide for the orthologs of *D. simulans* in the FlyBase database. Enabling the gridline option on the images in the ImageMaster software, the protein of interest was putatively located based on its theoretical pI and molecular weight. After identification, the spots were labeled and the images were imported including the overlaid labels and gridlines.

3. RESULTS

3.1 Gene expression in interspecies male hybrids:

Amplification of spermatogenesis candidate genes and two reference genes were based on primers designed by the alignment of conserved sequence regions between *D. simulans*, *D. sechellia* and *D. melanogaster*. It is possible that the primers might amplify a different region in the *D. mauritiana* samples due to mismatches. My sequencing of the amplified PCR products in *D. mauritiana* show high conservation with *D. simulans* sequences for all the targeted genes and reference genes. (Appendix XII). The sequences of primers show 100% conservation between *D. simulans* and *D. mauritiana* which is reflected by the fact that, except for *bgcn* (92%), all candidate gene and reference gene primers resulted in 100% efficiency.

An overall trend of under expression for all candidate genes were observed in the testes samples from interspecies hybrids compared to parental species (Figure 3.1 and 3.2). When using *RpL32* as a control gene for normalization, three genes of spermatogenesis showed a significantly lower level of expression in the sterile hybrids compared to their parents when using either ANOVA, pair comparisons using a nonparametric test (Mann-Whitney) or a randomization approach. They were *bam* and *bgcn*, two genes with a role in mitosis (Tukey post-hoc *P* values: $bam_{si\text{-}F1}$, $P = 0.0001$, $bam_{ma\text{-}F1}$, $P < 0.0001$; $bgcn_{si\text{-}F1}$, $P = 0.0007$, $bgcn_{ma\text{-}F1}$, $P = 0.0250$) as well as the meiotic control gene *sa* ($sa_{si\text{-}F1}$, $P < 0.0001$, $sa_{ma\text{-}F1}$, $P < 0.0001$) (Figure 3.1). The results are consistent when using *Act5C* for normalizations (Tukey post-hoc *P* values: $bam_{si\text{-}F1}$, $P = 0.0026$, $bam_{ma\text{-}F1}$, $P < 0.0001$; $bgcn_{si\text{-}F1}$, $P < 0.0001$, $bgcn_{ma\text{-}F1}$, $P = 0.0038$; $sa_{si\text{-}F1}$, $P <$

0.0001, $sa_{\text{ma-F1}}$, P= 0.0001) except that the other meiotic control gene assayed, *can*, also show significant down regulation in the sterile hybrid relative to their parents (Tukey post-hoc P values: $can_{\text{si-F1}}$, P= 0.032, $can_{\text{ma-F1}}$, P= 0.0089) (Figure 3.2). All other genes showed non-significant lower expression in the sterile hybrids relative to both parents or at least one parental species (dominance) regardless of what gene was used as control for normalizations (Figures 3.1 and 3.2).

Amplification plots showing Ct values of each reference gene by parental and hybrid samples found almost identical Ct values (*D. simulans* = 15.19, *D. maurtiana* = 15.51 and hybrid = 15.51 for *RpL32* and *D. simulans* = 15.5, *D. mauritiana* = 15.7 and hybrid = 15.72 for *Act5C*) which is noteworthy (Figure 3.3). Similar Ct values obtained for the reference genes, implies that there was an equal amount of total cDNA pool in the reaction mix. Therefore, normalized cDNA pools across samples can be used to report any misregulation that is gene specific and not due to low availability of initial template for amplification.

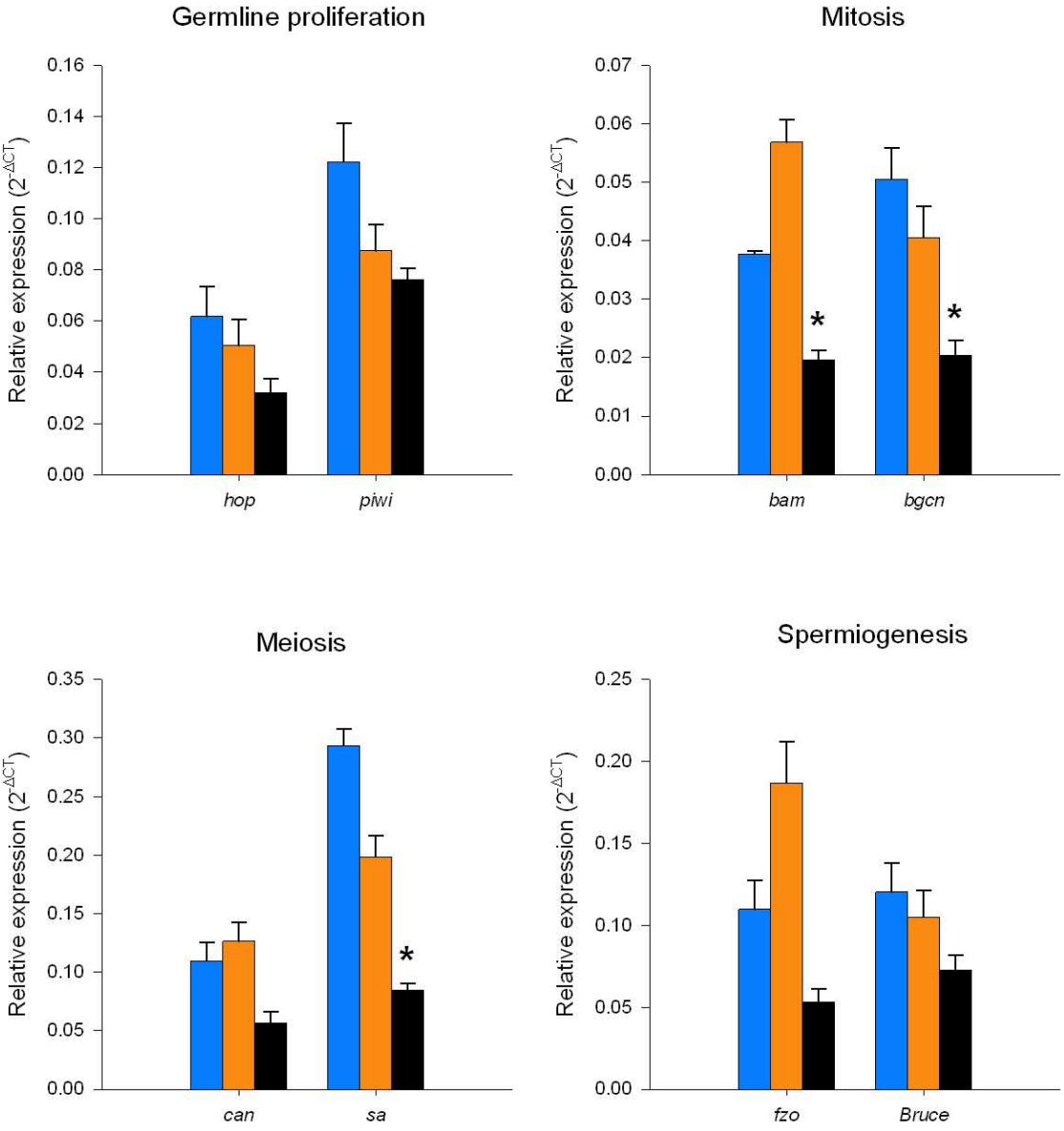

Figure 3.1. Relative average expression of candidate genes normalized with *RpL32* in parental and interspecies hybrid testes samples. Average expressions of *D. simulans* (blue bars), *D. mauritiana* (orange bars) and F1 hybrid (black bars) are plotted with standard errors. Genes that are significantly under expressed in hybrid in comparison to their parental expression are denoted with an asterisk

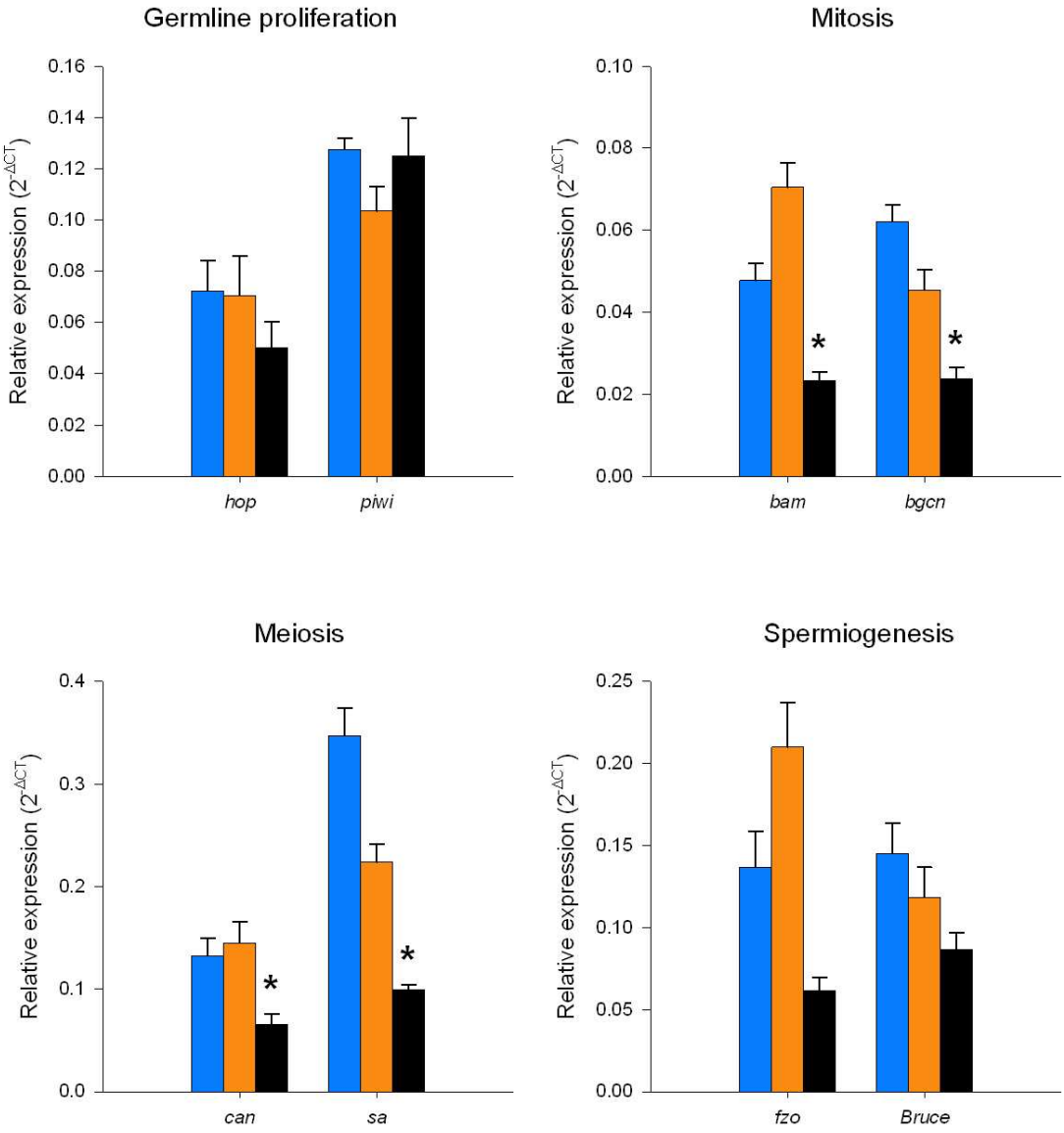

Figure 3.2. Relative average expression and standard error of candidate genes normalized with *Act5C* in parental and interspecies hybrid testes samples. The labels are as in figure 3.1.

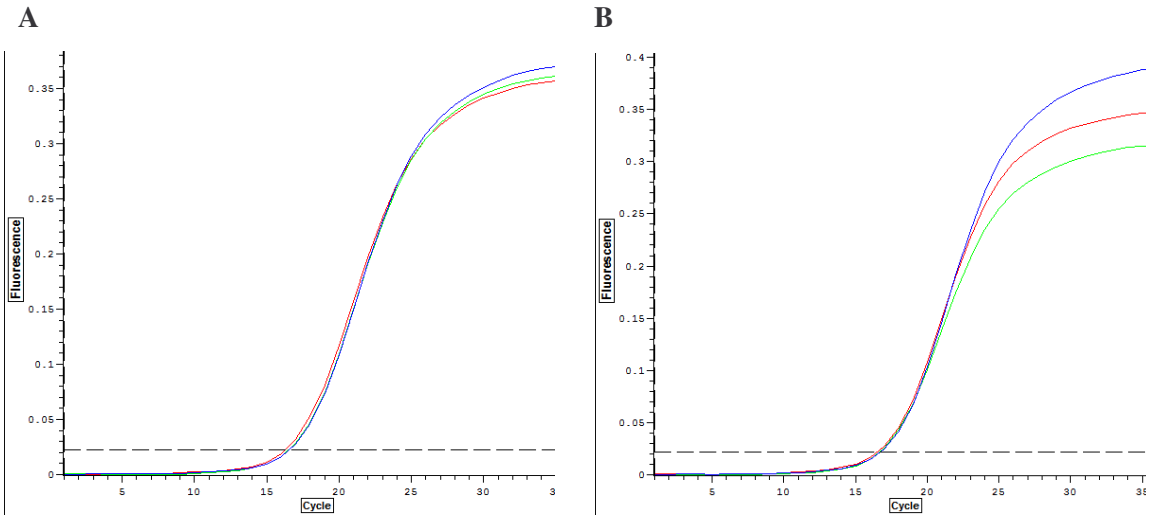

Figure 3.3. Amplification plots of reference genes by parental and interspecies hybrid samples. (A) *RpL32* and (B) *Act5C*. Each curve represents single sample; *D. simulans* in red, *D. mauritiana* in blue and F1 hybrid in green. Reaction cycle number is shown in the x-axis and the amount of fluorescence is shown on the y-axis.

A previous study found non-significant differences between parental *D. simulans* and *D. mauritiana* species and interspecies sterile male hybrids for premeiotic stage genes *bam* and *bgcn* (Moehring et al. 2007). It is possible that the use of RNA extractions from whole flies in the previous study might have made impossible to detect differences that are limited to the testes. Therefore the analysis of gene expression from parental and sterile hybrid males using RNA extraction from whole flies rather than testes was repeated. I found, in agreement with the previously reported results (Moehring et al. 2007) that all candidate genes showed a non-significant difference in expression between the parental and the hybrid flies, when whole body samples are used for RNA extractions (Figure 3.4).

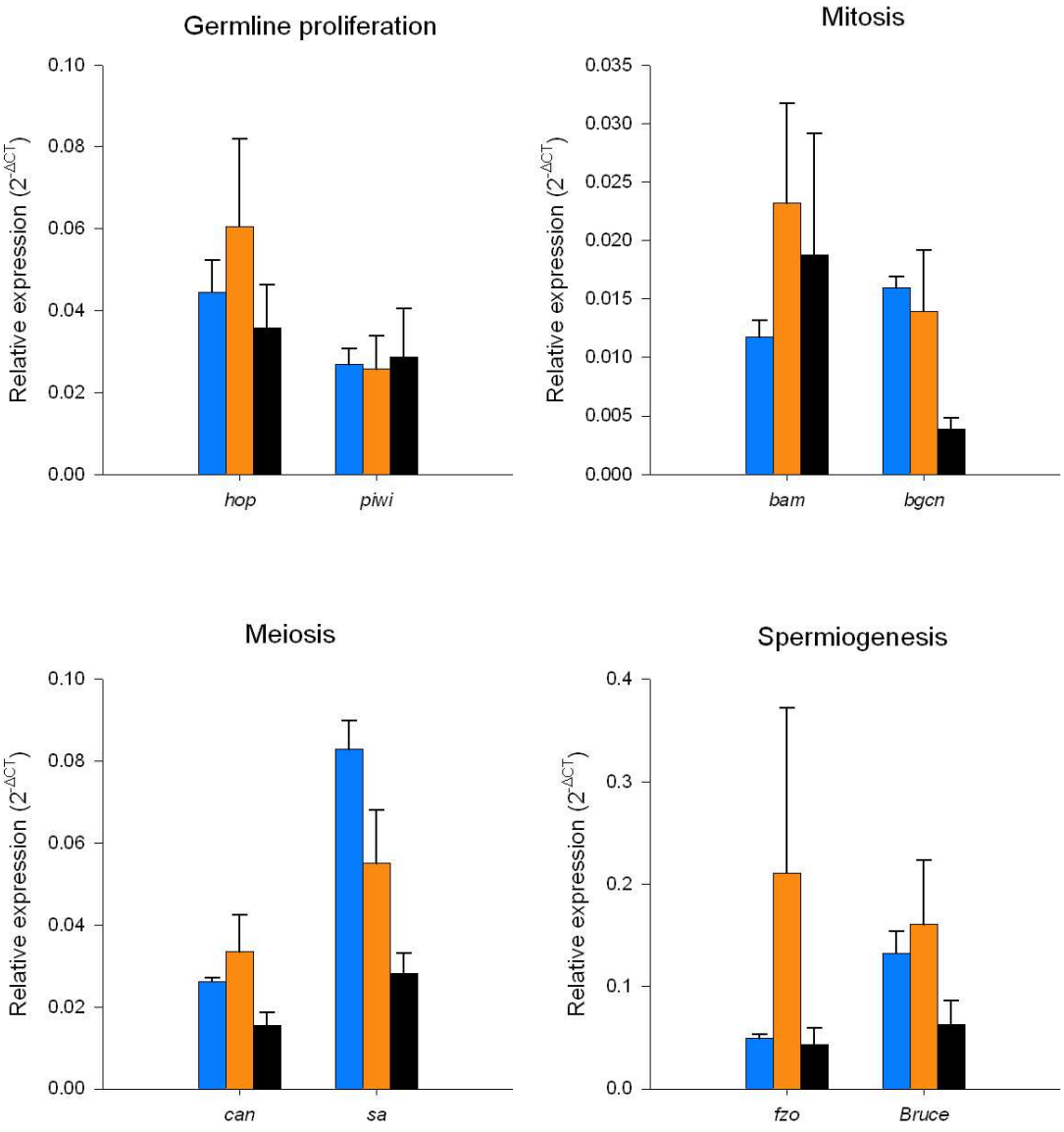

Figure 3.4. Relative average expression and standard error of candidate genes in parental and interspecies hybrid whole body samples. The average expression of candidate genes are normalized with *RpL32*The labels are as in figure 3.1.

3.2 Gene expression in interspecies female hybrids:

Gene down regulation could be due to overall failures in hybrids that would affect males and females rather than being specific to only males. Data from *D. melanogaster* shows that hopscotch (*hop*) is a JAK tyrosine kinase involved in the JAK-STAT signaling pathway and its expression is detected in the follicular epithelium to control follicle cell differentiation during oogenesis (Silver and Montell 2001, McGregor et al. 2002). *piwi* is required for germline stem cell renewal of both the sexes and expressed in the female germarium (Cox et al. 2000, Wong et al. 2005). Interaction of *bam* and *bgcn* is necessary for cystoblast cell differentiation and *bgcn* is expressed at very low levels in oogenesis (Gonczy et al. 1997, Ohlstein et al. 2000). In germaria and mid-stage egg chambers, *Bruce* is expressed and acts to inhibit autophagy and cell death (Hou et al. 2008).

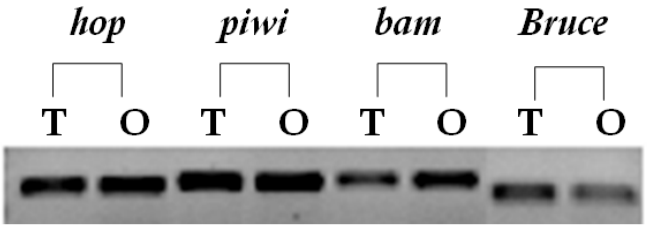

Figure 3.5. PCR amplification of cDNA samples of candidate genes in *D. simulans* testes (T) and ovaries (O) samples.

My PCR results show that *hop*, *piwi, bam* and *Bruce* are similarly expressed in *D. simulans* ovaries and testes (Figure 3.5). For genes that are equally expressed in ovaries and testes (*hop, piwi, bam*, *Bruce*), it is possible to test whether their under expression in sterile hybrid males is simply a consequence of misregulation of gene expression in hybrid flies. Gene expression analysis through qRT-PCR was performed from both

parental and fertile hybrid ovary samples from interspecies crosses (*D. simulans* × *D. mauritiana*). The results show no significant differences in gene expression for any of the genes tested regardless of what gene was used as control and of what statistical approach was used to compare the samples (Figure 3.6). This result confirms that the down regulation observed in sterile male hybrids is specific to the males and not a general consequence of hybrid dysfunction.

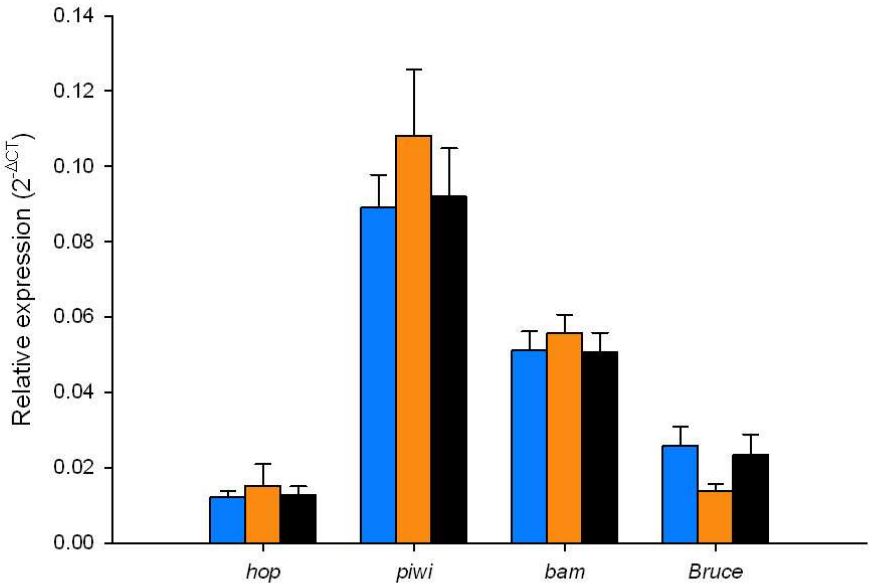

Figure 3.6. Relative expression and standard error of candidate genes normalized with *RpL32* in parental and hybrid ovaries samples. The labels are as in figure 3.1.

3.3 Gene expression in intraspecies and interspecies hybrids:

A significantly lower level of gene expression between parental species and sterile interspecies hybrids could be linked to interspecies divergence in regulatory elements and a breakdown in gene regulation during spermatogenesis. If so, intraspecific hybrids between populations that produce fully fertile hybrids should not show

significant drops in gene expression. This hypothesis could be tested by assaying gene expression differences between flies from two parental *D. simulans* populations (*D. simulans* California and *D. simulans* Congo) and their intraspecific male hybrids. The results show no significant differences in gene expression between both parental strains and the intraspecific hybrids with the results being consistent regardless of which gene was used as control for normalizations (Figures 3.7 and 3.8). The intraspecific result serves as a control that further supports that the down regulation in sterile interspecies male hybrids is linked to gene regulatory or developmental problems in the sterile hybrid.

The consistent result of down regulation found in sterile hybrids relative to their parents, regardless of what gene is used as control for normalization, for *bam*, *bgcn* and *sa* suggests that down regulation at these genes or upstream genes in their pathway is a possible direct cause for hybrid male sterility. However, if this is the case, the level of gene expression in sterile hybrids should be significantly lower than both parents as well as any other strain of either *D. simulans* and *D. mauritiana*. The above results show that this is true only for *bam* and *sa* when using *RpL32* as a control gene but the results do not hold when *Act5C* is used to normalized the data (Figures 3.7 and 3.8).

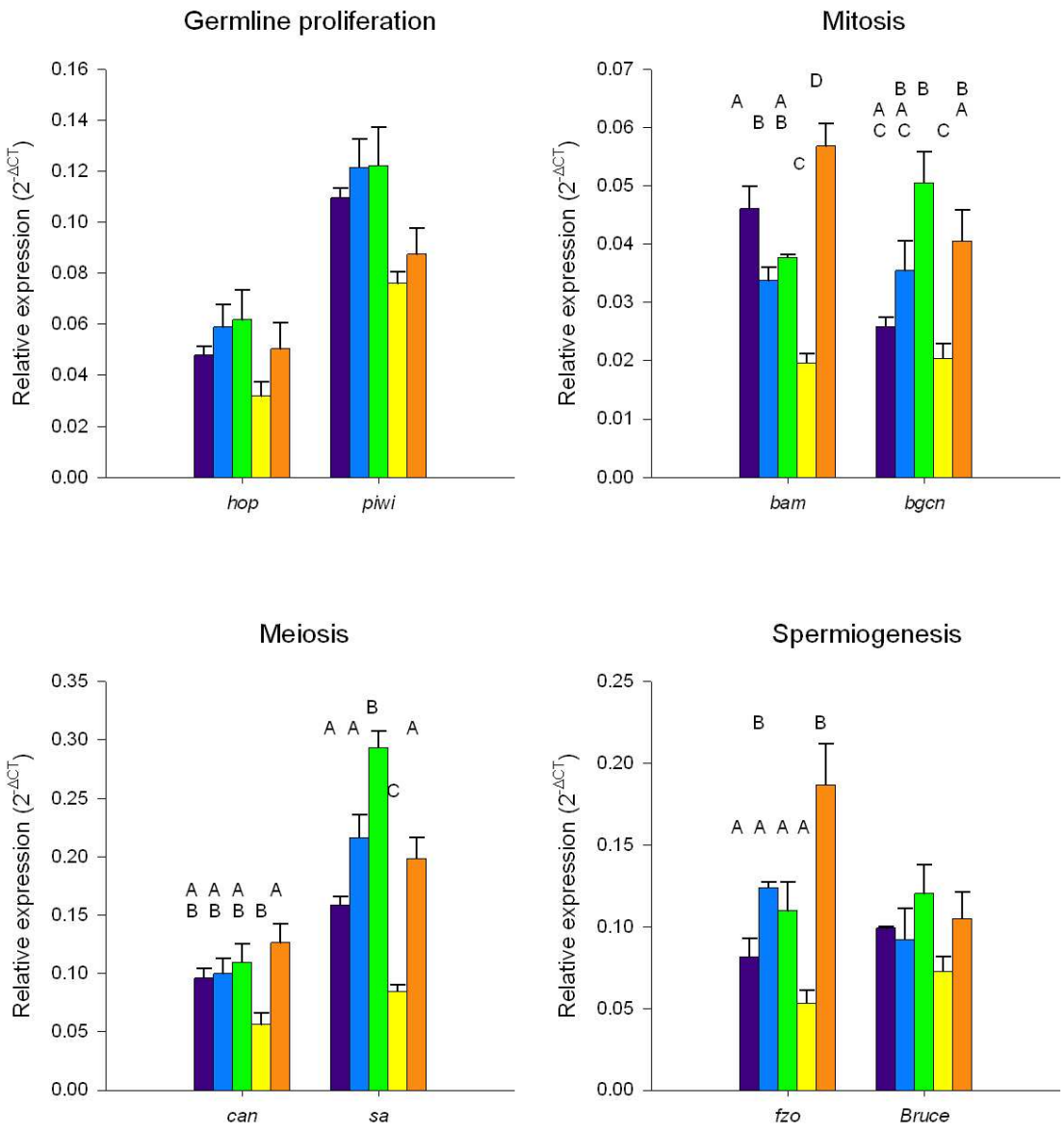

Figure 3.7. Relative expression and standard error candidate genes normalized with *RpL32* from parental and hybrid (intra- and inter-) testes samples. Average expression are plotted for *D. simulans* Congo (purple bars), intraspecies (*simulans*) fertile hybrid (blue bars), *D. simulans* California (green bars), sterile interspecies (*simulans/mauritiana*) hybrids (yellow bars) and *D. mauritiana* (orange bars). Shared letters above columns indicate that averages are not statistically different.

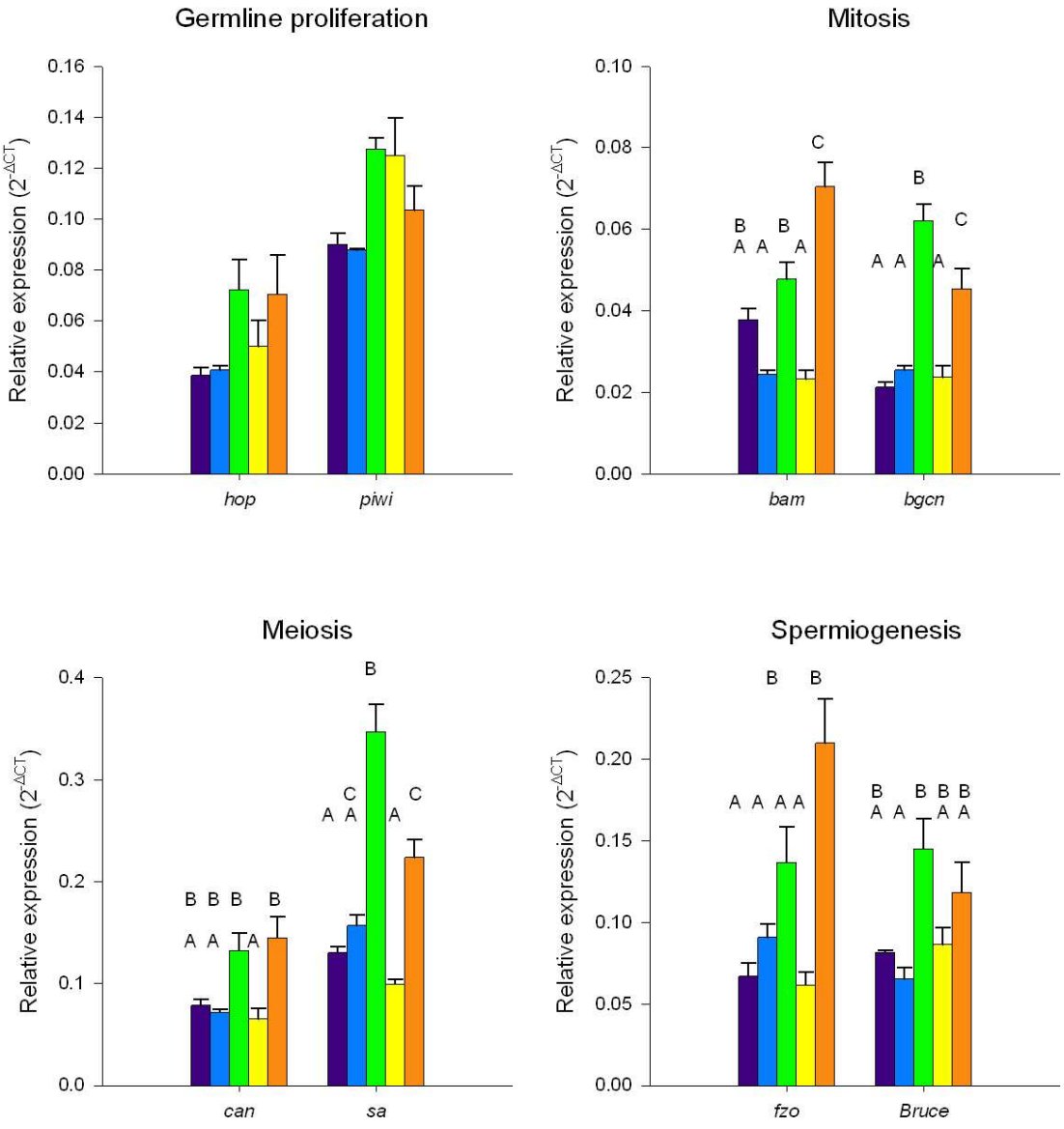

Figure 3.8. Relative expression of candidate genes normalized with *Act5C* from parental and hybrid (intra- and inter-) testes samples. Labels are as in figure 3.7.

3.4 Protein expression in parental species and hybrids:

Total protein extraction from parental (*D. simulans* and *D. mauritiana*) and F1 hybrids revealed that good quality protein extraction could be used for further separation of proteins using a two-dimensional system (Figure 3.9).

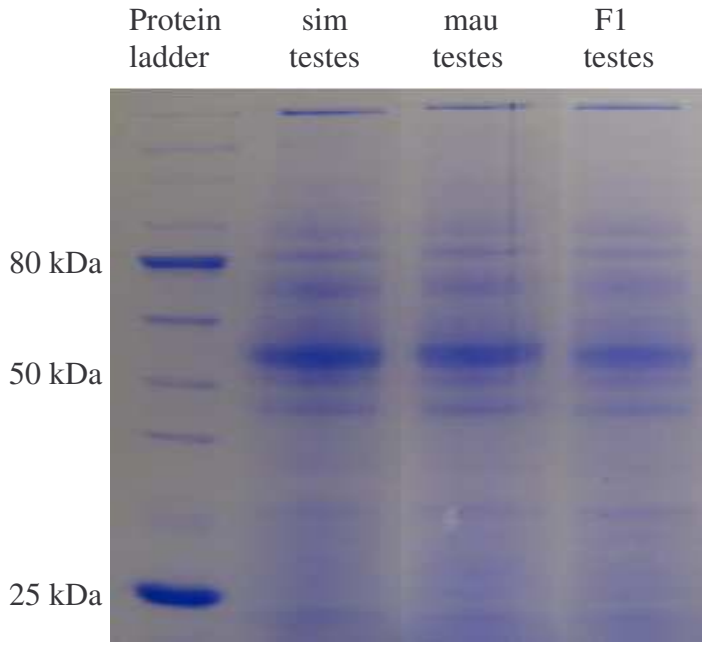

Figure 3.9. SDS gel electrophoresis from parental and interspecies hybrid testes samples. *D. simulans* (sim), *D. mauritiana* (mau) and F1 hybrid (F1) showing abundance of protein in their respective extractions.

After repeated trials and trouble shooting, the second dimension electrophoresis of the parental and the hybrid samples yielded a streaked pattern of protein spots, which were not distinctive. Although all three samples (*D. simulans* (cal), *D. mauritiana* and F1 hybrid) were run on the same second dimension SDS gel, *D. mauritiana* and F1 hybrid samples yielded fairly resolved protein spots with minimal streaking. However, *D. simulans* yielded poor quality spots with heavy streaking and were excluded from further analysis. The total number of spots identified in *D. mauritiana* and F1 hybrid by the

ImageMaster 2D platinum software was 131 and 162 respectively. By matching both the gel images, 104 spots (79.3%) were identified as shared between the parental and hybrid gels. Protein spots within the molecular weight range of 66 kDa to 29 kDa showed a better resolution with the vector patterns showing maximal consistencies. Within this region, protein products of the candidate genes *bag-of-marbles (bam)* and *spermatocyte arrest (sa)* were putatively identified in *D. mauritiana* and F1 hybrid gel images, based on their theoretical pI and molecular weight (Figures 3.10, 3.11). Preliminary quantification of the protein spots using ImageMaster 2D platinum showed little difference in volume or intensity between the *D. mauritiana* and F1 hybrid for *bam* (Table 3.1). A difference was observed in the spot volume and intensity for *sa* between the F1 hybrid and *D. mauritiana*, with the F1 hybrid showing higher values than the parental spot.

Table 3.1. Physico-chemical properties of *bam* and *sa*. The theoretical pI and molecular weights was determined using Protparam tool of Expasy. Volume and spot intensity was calculated using ImageMaster2D platinum.

Candidate gene	**pI**	**Molecular weight**	**Volume**		**Spot intensity**	
			D. mauritiana	**F1 hybrid**	***D. mauritiana***	**F1 hybrid**
bag-of-marbles (bam)	6.21	50 kDa	9034.45	9364.51	663	689
spermatocyte arrest (sa)	5.1	31 kDa	1890.8	4585.95	204	341

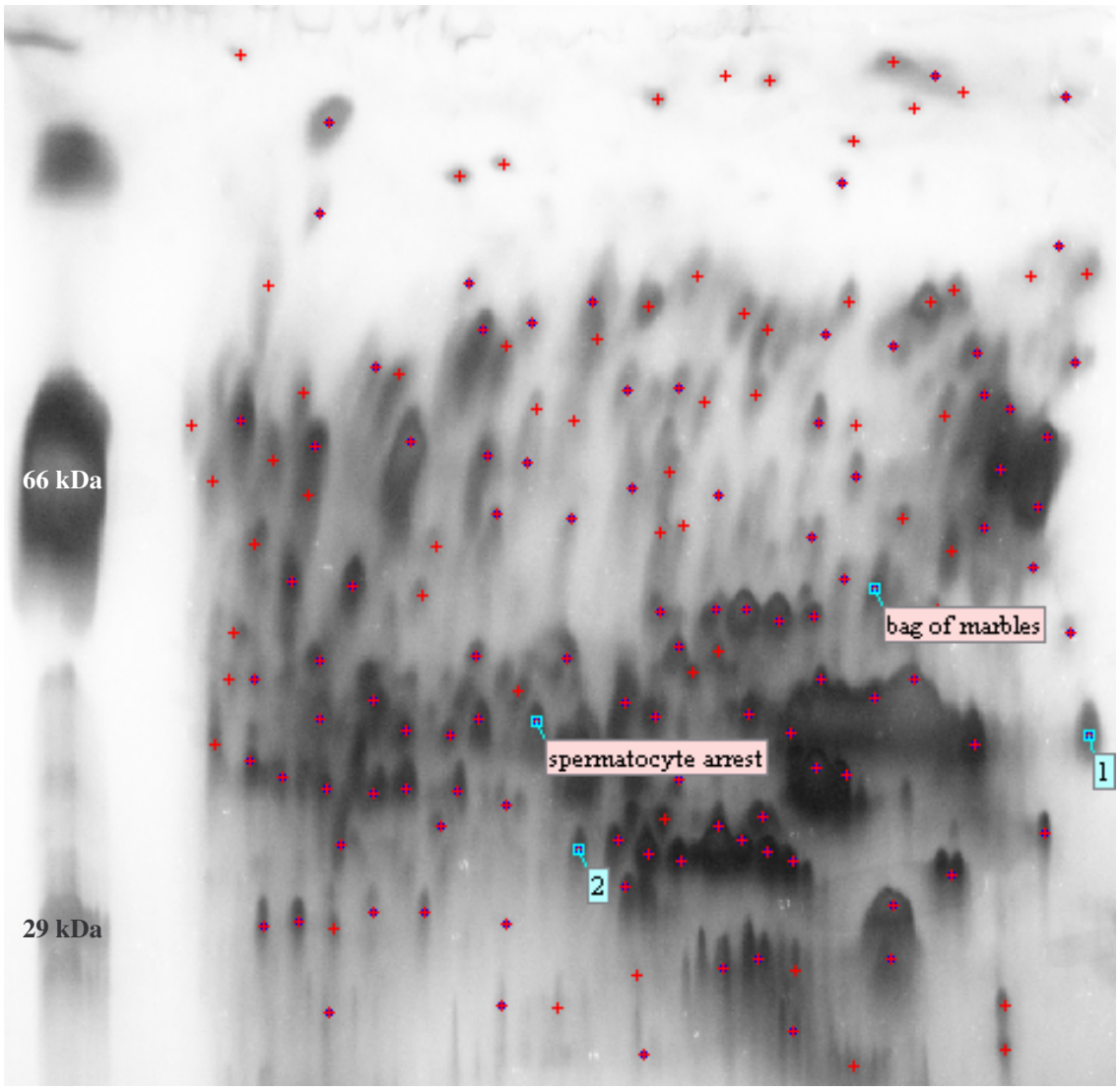

Figure 3.10. Second-dimensional SDS image of F1 hybrid testes samples. *bag-of-marbles (bam)* and *spermatocyte arrest (sa)* are labeled on the gel and two spots designated as landmarks are labeled as 1 and 2. Molecular marker with molecular weight 66 kDa and 29 kDa are marked on the gel.

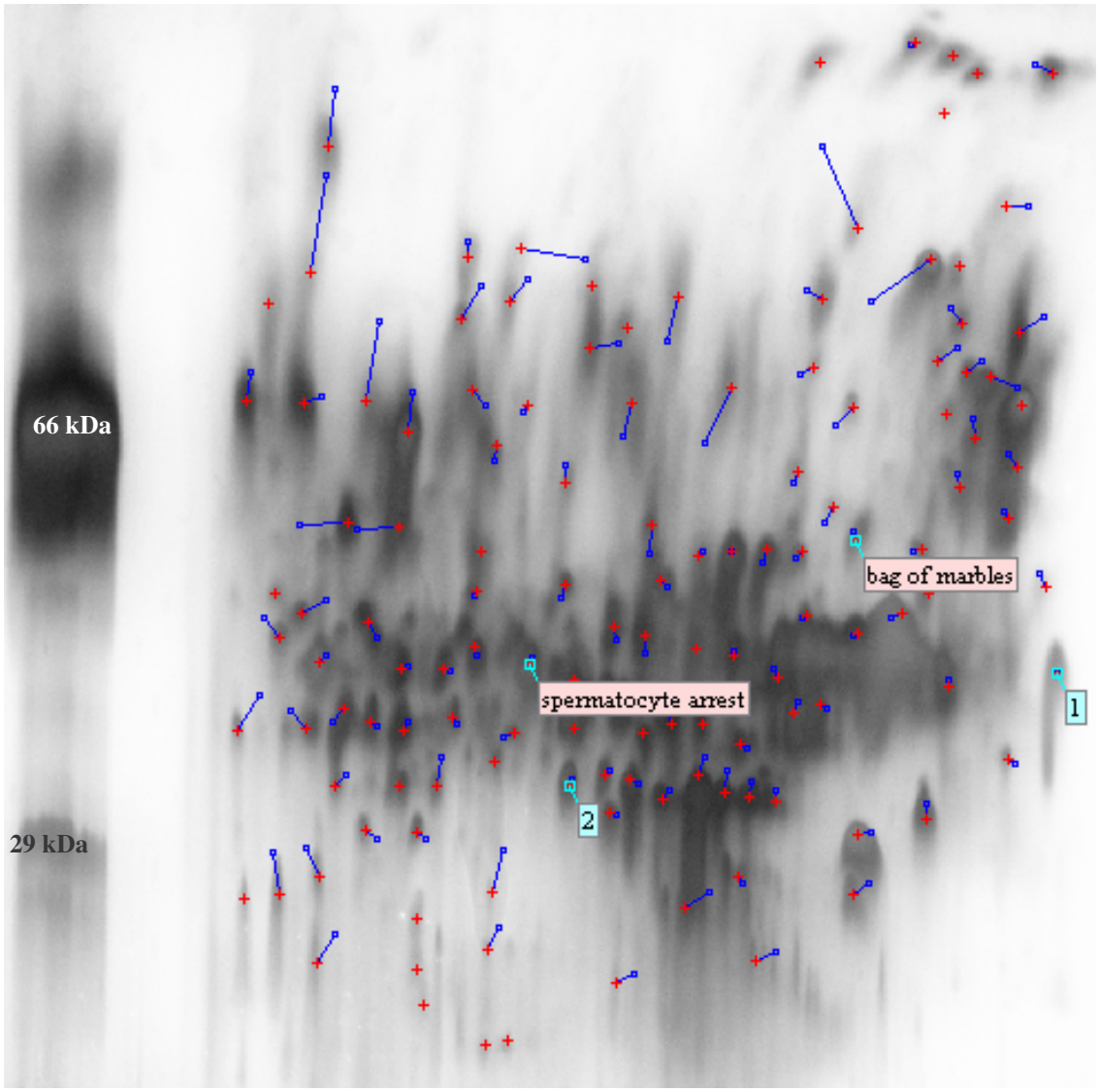

Figure 3.11. Second-dimensional SDS image of *D. mauritiana* testes samples. Labels are as in figure 3.10.

3.5 Testes morphology in interspecies hybrids:

Although interspecies hybrid males generated between *D. simulans* (female) and *D. mauritiana* (males) are always sterile, the morphology of the testes was previously reported to be completely intact and similar to wild type fertile flies except that the seminal vesicle lacks sperm (postmeiotic arrest; Lachaise et al. 1986, Perez 1993, Kulathinal and Singh 1998). During dissections, I observed 1027 pairs of hybrid testes and found that 996 showed morphology similar to the parental species (Figure 3.12 b). However, 31 testes showed disruption in their apical region in one or both of the testes isolated from F1 hybrid flies (Figure 3.12 c-f). Since the apical region is where the germline proliferation and mitosis developmental stages occur, observation of these phenotypes could be correlated with the possibility of premitotic disruption at the cellular level. I confirmed, by puncturing the seminal vesicle, that all interspecies hybrid testes did not have any sperm.

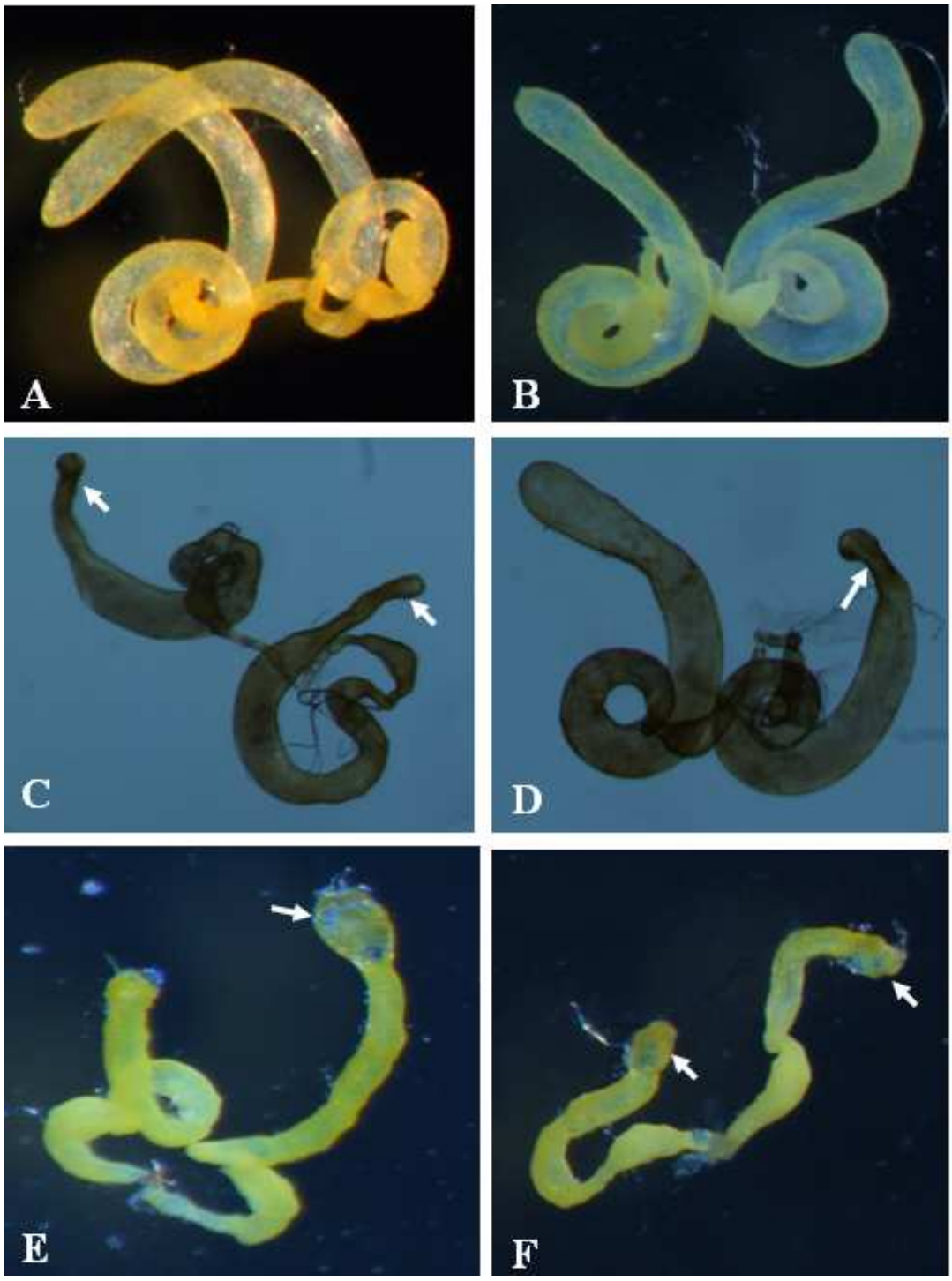

Figure 3.12. Morphology of parental and interspecies hybrid testes. The testes were dissected from 4-6 day old flies and observed under a stereoscopic zoom microscope at 70-90X. (A) Testes from *D. simulans*. (B) Testes from F1 hybrid showing no difference in morphology when compared to *D. simulans*. (C-F) Testes from F1 hybrid showing atrophy when compared to *D. simulans*. Arrow heads show the atrophied apical region of the testes.

DISCUSSION

Along with several previous studies on the genetic architecture of hybrid male sterility, particularly those in the *D. simulans* clade (Palopoli and Wu 1994; Perez and Wu 1995; Tao et al. 2003a, 2003b), my results support a polygenic view for hybrid male sterility leading to developmental disruption in hybrids. In particular, the results of this research suggest a non stage-specific transcriptional down regulation of spermatogenesis genes exclusive to interspecific sterile male hybrids. The following reasons could explain that this investigation is distinctive from previous studies. Firstly, I surveyed genes from all developmental/transition stages in spermatogenesis i.e. from germline proliferation to spermiogenesis. Secondly, the expression differences were analyzed at the very site of hybrid disruption i.e. testes, isolated from adult parental and hybrid flies. This strategy helped to identify the contribution of differences in gene expression towards hybrid male sterility. Thirdly, the selection of genes essential to the *Drosophila* oogenesis pathway (*hop*, *piwi*, *bam* and *Bruce*), enabled the examination of whether differences in the expression of parental and hybrids genes was male-specific rather than hybrid status. Lastly utilization of quantitative RT-PCR to validate the differential expression between species and hybrids, precluded the possibility that the differences in expression were due to hybridization bias, as observed in previous studies when microarrays were employed (Michalak and Noor 2003; Haerty and Singh 2006; Moehring et al. 2007).

My gene expression assay results show that all candidate genes have relatively lower average expression in the interspecific sterile hybrid testes than their parental species, irrespective of the gene's role and stage of the spermatogenesis pathway. The

gene expression pattern that is observed in sterile interspecies hybrid testes samples appears at first sight to be in discrepancy with previous results (Moehring et al. 2007; Catron and Noor 2008). Moehring et al. (2007) utilized RNA isolated from whole body samples and identified a non-significant down regulation of *bam*, *bgcn*, *can* and *fzo* in the sterile hybrids compared to parental samples. The disagreement between their results and those presented here could be related to the source of RNA sample that was employed to assay gene expression. Concentration of gene products could be accurately determined at the site of synthesis (i.e. testes in this study) whereas it is likely to be attenuated when measured in whole body samples. Indeed, my results from whole body extractions also show no significant differences for *bam*, *bgcn*, *can* and *fzo* which is in agreement with results by Moehring et al. (2007). Selection of an entirely different class of meiotic (*bam* and *bgcn*) and post-meiotic genes (*can* and *sa*) in this study could explain the disparity with results by Catron and Noor (2008). They reported that F1 hybrids are more likely to display under expression of postmeiotic genes (*Mst84D* and *donjuan*) than premeiotic genes (*aly* and *comr*). Transcriptional and translational regulation of *aly* in mutant males have been shown to be different and independent of other meiotic arrest genes like *mia*, *can*, *sa* (Lin et al. 1996; White-Cooper et al. 2000). The results could possibly indicate that *can*-class of genes (*can*, *mia*, *sa*, *rye*) are subjected to down regulation in sterile hybrid spermatogenesis, while *aly*-class of genes (*aly*, *comr*, *achi/vis*, *topi*) might not. However, this explanation will leave unresolved why mitotic division control genes such as *bam* and *bgcn* are subjected to significant down regulation in this study. A possible explanation might relate to the independent nature of the spermatogenesis pathway, showing that mutations in spermatogenesis genes usually cause cytological defects in one

or few developmental events, but the overall developmental pathway can still proceed. While mutations of *bam* and *bgcn* have been shown to arrest the entire process of spermatogenesis (Fuller 1998), it is unclear whether differences in expression can be circumvented by other classes of genes in a way similar to meiotic control genes. The expression results from intraspecific testes samples did not show significant differences between both parents and hybrid samples, denoting that any significant down regulation of spermatogenesis genes is related to interspecific divergence in male regulatory control. It could be also explained due to simple allometric changes in sterile hybrid males developing abnormally in terms of cell types and mRNA abundance relative to parents.

The results imply discordance between differences in gene expression (no stage-specific down regulation) and phenotypic changes in sterile hybrids (postmeiotic arrest). This discrepancy could be explained by several hypotheses: One possible explanation is a major transcriptional shut down before meiosis during spermatogenesis, with a very limited post-meiotic transcription. Most transcripts required throughout spermatogenesis are transcribed prior to meiotic divisions (Oliveri and Oliveri 1965; Gould-Somero and Holland 1974; Barreau et al. 2008), so it is possible that during spermatogenesis in sterile hybrids, most transcripts could be down regulated as soon as they are synthesized. If this were to happen, a general down regulation of genes would be observed regardless of the gene's role within the developmental pathway. Shared down regulation of overall transcripts in sterile hybrids might be linked to the presence of intercytoplasmic connection between developing spermatocytes. During *Drosophila* spermatogenesis, interconnected cytoplasmic bridges that initially appear during the first mitotic division (due to incomplete cytokinesis) persist into the later stages of spermatid elongation

(Fuller 1993, Hime et al. 1996). Earlier hypotheses to explain the requirement of intercellular bridges during mouse spermatogenesis have discussed that these bridges play a crucial role in the distribution of equal amounts mRNAs and cytoplasmic communication between the developing germ cells (Morales et al. 1998; Hecht 2000; Ventela et al. 2003). From the results of overall down regulation in transcription, it is possible that low amounts of available transcripts might be distributed through these cytoplasmic bridges. This could contribute to a further spread of gene expression down regulation among the developing cells as well as to progress towards the subsequent stage in the pathway. Alternatively, the lower amount of transcript abundance observed throughout hybrid spermatogenesis and the observation of postmeiotic arrests in *D. simulans* females × *D. mauritiana* males sterile hybrids (Lachaise et al. 1986; Wu et al. 1992; Kulathinal and Singh 1998) could be explained if a threshold of transcripts and protein products are needed for cellular progression into subsequent stages of spermatogenesis. A threshold phenotype might explain why spermatogenesis progresses in sterile hybrids through mitosis with a final "built up" breakdown at a postmeiotic stage. In fact, microscopic observation of sterile hybrid testes from my results and from earlier studies reveal that threshold at the cellular level appears to be somehow flexible as occasional premeiotic arrest occur between species of the *D. simulans* clade (Lachaise et al. 1986; Zeng and Singh 1993; Kulathinal and Singh 1998). It is also worth noting that detailed cellular studies of hybrid male sterile testes have not been conducted and so the general observation of presence of sperm bundles in a sterile male might not necessarily mean that subtle premeiotic problems might have occurred during development. Until further cellular and functional assays are performed, it is hard to conclude with certainty

whether down regulation of gene expression in testes of sterile hybrid males might have any effect on the sterility phenotype or if it is merely a consequence of sterility.

Verification of expression pattern in hybrid oogenesis helped to establish whether the observed down regulation is exclusive to interspecies hybrid, or a general down regulation pattern in all hybrid progenies. Hybrid females in this particular cross are fertile offspring, so I expected no difference in expression between the parental and hybrid samples. The expression results of *hop*, *piwi*, *bam* and *Bruce* did not show any siginificant difference in expression between parental and hybrid samples, suggesting that significant down regulation of spermatogenesis genes could be correlated with hybrid male sterility and not a general pattern to be observed in interspecific progenies. Examination of the Ct values obtained for the reference genes and all candidate genes exemplify three major points: 1. Very similar amplification curves by the parental and hybrid samples of the reference genes implies that similar amount of total cDNA were used in the reaction mixes. 2. Amplification differences observed in the candidate genes (earlier amplification of the parental and later amplification of the hybrid) show that only these genes are subjected to regulation while the reference genes are not. 3. Normalized cDNA pools across samples will precisely report any misregulation that is gene specific and not due to low availability of initial template for amplification.

While interspecies sterile hybrids show an exclusive down regulation at the transcriptional level, analysis of protein expression using a 2D electrophoresis system did not detect down regulation in the F1 hybrid compared to the parental species. A similar investigation between two sibling species of the *D. simulans* clade (*D. simulans* and *D.sechellia*) to detect the number of sterility genes reported that 97.4% of the species-

specific protein spots (89 in *D. simulans* and 79 in *D. sechellia*) were found in the interspecies hybrids and at a level of expression that were intermediate in between the two parental species (Zeng and Singh 1993). My protein expression analysis between *D. simulans* and *D. mauritiana* pairs identified that the sterile hybrids share 79.3% of proteins with *D. mauritiana* testes samples. The exclusion of a protein profile of one of the parental species (*D. simulans*) due to poor resolution precludes me from determining what proportion of the remaining 20.7% of spots in *D. mauritiana* are species specific. The use of a much narrower pH range (4-7) in this study compared to that used by Zeng and Singh (1993) (pH 3-10) might explain a prominent reduction in the total number of spots identified in this thesis (131 spots in *D. mauritiana* and 162 spots in F1 hybrid) relative to the previous study (1063 spots in *D. simulans* and 983 spots in *D. sechellia*). Moreover, the goal of my protein investigation was to identify whether gene products of the candidate genes which are subjected to significant transcriptional down regulation, show any misregulation at the translation level. Putative identification of the gene product of *bam* did not show qualitative differences in expression between the parental and the hybrid samples. In the case of *sa*, the spot observed in the hybrid showed a much higher value in terms of spot volume and intensity which could be due to the close migration of the two proteins with similar physico-chemical properties detected as a single spot. The fact that both the protein spots were identified in the F1 hybrid suggests that my results are in agreement with the results by Zeng and Singh (1993) that the interspecific hybrid shows equal protein expression when compared to the parental species. It is also noteworthy from my results that a significant reduction in the level of transcription might not be drastic enough to affect the protein synthesis. This seems to be

in correspondence with my gene expression assay results which show that regardless of lower transcript abundance of premeiotic genes, spermatogenesis in sterile hybrids is likely to progress until after meiosis to generate a final "built up" breakdown. However, caution should be exercised in the interpretation of the protein expression part of the results because of substandard resolution of all the gels and one of the parental species having to be excluded for any further analysis.

To explain the overall transcriptional down regulation observed for genes of spermatogenesis and the significant drop for genes that control key transitions, it is necessary to consider results from prior studies showing a widespread pattern of rapid evolution of sex-related genes in *Drosophila* at both coding and gene expression levels (Civetta and Singh 1998, Haerty et al. 2007). An analysis of the same group of genes in terms of expression and coding sequence divergence have shown under expressed genes within the testes of interspecies hybrids to evolve more rapidly at the protein sequence level than nonmisregulated genes or overexpressed misregulated genes (Artieri et al. 2007). Moreover, molecular population studies have shown rapid divergent coding sequence evolution driven by positive selection between closely related species of *Drosophila* for two of the genes assayed in this study (*bam* and *bgcn*) (Civetta et al. 2006; Bauer-Dumont et al. 2007). Results from this study can be broadly understood in the context of rapid divergence of male regulatory elements and the male sex drive theory of evolution which has suggested that the male driven effects of selection arising through the actions performed by males can lead to a masculinization of the genome with increased interspecies divergence and increased turnover of male-expressed genes (Singh and Kulathinal 2005). However, due to the lack of detailed cellular analysis of

developmental problems in sterile male hybrids testes, it is hard to rule out whether gene down regulation is simply an allometric byproduct of development.

While I suspect that an allometric cellular byproduct will not differentially affect some genes more drastically than others as observed in my results, the contrasting hypothesis could be tested in future studies by either analysis of gene expression in fully fertile backcross males with different *D. mauritiana* introgressions in an isogenic *D. simulans* background. Whether divergence in cis- or trans-regulatory elements contribute to the down regulation of genes of spermatogenesis in hybrids could be tested using fertile introgressed male progenies that are partial hybrids for different chromosomes in an otherwise isogenic *D. simulans* background.

CONCLUSION

Through genomic and proteomic approaches to decipher hybrid male sterility in *Drosophila* species, this study has identified the following findings:

1. An overall trend of lower average expression of genes in the interspecies sterile *Drosophila* hybrids from different stages of sperm development with significant down regulation at specific mitotic and meiotic control genes.
2. Down regulation is only evident in interspecies testes expression at the transcriptional level, clearly indicating that the problem is male-specific and exclusive to gene expression, not to protein expression.
3. Significant lower hybrid testes expression is not necessarily detected when compared to a non-parental strain of *D. simulans*, ruling out these genes as directly responsible for the hybrid sterile phenotype.
4. Down regulation in the interspecies hybrid is the result of rapid evolution at male-specific regulatory elements or simply a byproduct of subtle cellular problems during development.

References:

Allison DB, Cui X, Page GP, Sabripour M (2006) Microarray data analysis: from disarray to consolidation and consensus. Nat Rev Genet 7:55-65.

Arama E, Agapite J and Steller H (2003) Caspase activity and a specific cytochrome C are required for sperm differentiation in *Drosophila*. Developmental Cell 4: 687-697.

Artieri CG, Haerty W and Singh RS (2009) Ontogeny and phylogeny: molecular signatures of selection, constraint, and temporal pleiotropy in the development of *Drosophila* BMC Biology 7: 42.

Barreau C, Benson E, Cooper HW (2008) Comet and cup genes in *Drosophila* spermatogenesis: the first demonstration of post-meiotic transcription. Biochem. Soc. Trans. 36: 540-542.

Bauer DuMont VL, Flores HA, Wright MH, Aquadro CF (2007) Recurrent positive selection at Bgcn, a key determinant of germ line differentiation, does not appear to be driven by simple coevolution with its partner protein Bam. Mol Biol Evol 24:182-191

Bock IR (1984). Interspecific hybridization in the genus *Drosophila.* Evol. Biol. 18: 41-70.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248-254

Cabot EL, Davis AW, Johnson NA and Wu C-I (1994) Genetics of reproductive isolation in the *Drosophila simulans* clade: Complex epistasis underlying hybrid male sterility. Genetics 137: 175-189.

Cagan RL (2003) Spermatogenesis: Borrowing the apoptotic machinery. Current Biology 13: R600-R602

Catron DJ and Noor MAF (2008) Gene expression disruptions of organism versus organ in *Drosophila* species hybrids PLoS ONE 3: e3009.

Cazemajor M, Joly D and Montchamp-Moreau (2000) Sex-ratio meiotic drive in *Drosophila simulans* is related to equational nondisjunction of the Y chromosome. Genetics 154: 229-236.

Charlesworth B, Coyne JA, Barton NH (1987) The relative rates of evolution of sex chromosomes and autosomes. Am. Nat. 130: 113-146.

Civetta A and Singh RS (1995) High divergence of reproductive tract proteins and their association with postzygotic reproductive isolation in *Drosophila melanogaster* and *Drosophila virilis* group species. J Mol. Evol. 41: 1085-1095.

Civetta A, Singh RS (1998) Sex-related genes, directional sexual selection, and speciation. Mol Biol Evol 15:901-909

Civetta A, Rajakumar SA, Brouwers B, Bacik JP (2006) Rapid evolution and gene-specific patterns of selection for three genes of spermatogenesis in Drosophila. Mol Biol Evol 23:655-662

Cooper KW (1965) Normal spermatogenesis in *Drosophila*, in Biology of *Drosophila* (Demerec M) pp 1-61. Hafner publsishing company, New York.

Coulthart and Singh (1988a) High level of divergence of male-reproductive-tract proteins, between *Drosophila melanogaster* and its sibling species, *D. simulans*. Mol. Biol. Evol 5: 182-191.

Coulthart and Singh (1988b) Low genic variation in male reproductive tract proteins of *Drosophila melanogaster* and *D. simulans*. Mol. Biol. Evol 5: 167-181.

Coyne JA (1984) Genetic basis of male sterility in hybrids between two closely related species of *Drosophila*. Proc. Natl. Acad. Sci. 81: 4444-4447.

Coyne JA (1985) The genetic basis of Haldane's rule. Nature 314: 736-738

Coyne JA and Charlesworth B (1986) Location of X-linked factor causing sterility in male hybrids of *Drosophila simulans* and *D. mauritiana*. Heredity 57: 243-246.

Coyne JA and Orr HA (1989) Patterns of speciation in *Drosophila*. Evolution 43: 362-381.

Coyne JA, Charlesworth B and Orr HA (1991) Haldane's rule revisited. Evolution 45: 1710-1714.

Coyne JA (1992) Genetics and Speciation. Nature 355: 511-515.

Coyne JA and Orr HA (2004) The genetics of postzygotic isolation, in Speciation pp 283-319. Sinauer Associates, Inc, Sunderland, Massachusetts.

Cox DN, Chao A and Lin H (2000) *piwi* encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. Development 127: 503-514

Craft WA (1938) The sex ratio in mules and other hybrid mammals. Q. Rev. Biol. 13: 19-40.

Dobzhansky Th (1933) On the sterility of the interracial hybrids in *Drosophila pseudoobscura*. Proc Natl Acad Sci U S A. 19: 397–403.

Dobzhansky Th (1934) Studies on hybrid sterility. I. Spermatogenesis in pure and hybrid *Drosophila pseudoobscura*. 2. Zellf. mikr. Anat. 21: 169-223.

Dobzhansky Th (1970) Reproductive isolation, in *Genetics of the evolutionary process*. pp 311-350. Columbia University Press., New York and London.

Dobzhansky Th and Beadle GW (1936) Studies on hybrid sterility IV. Transplanted testes in *Drosophila pseudoobscura*. Genetics 21: 832-840.

Eberhard WG (2001) Species-specific genitalic copulatory courtship in sepsid flies (Diptera, Sepsidae, Microsepsis) and theories of genitalic evolution. Evolution 55: 93-102.

Frank SA (1991) Divergence of meiotic drive-suppression systems as an explanation for sex-biased hybrid sterility and inviability. Evolution 45: 262-267.

Freeman S and Herron JC (2001) Mechanisms of speciation, in Evolutionary analysis. Second edition pp 403-435. Prentice-Hall, Inc., New Jersey.

Fuller MT (1993) Spermatogenesis, in The Development of *Drosophila melanogaster* Vol.1 (Bate M, Martinez-Arias A, eds) pp71-147. Cold Spring Harbor Press, Cold Spring Harbor, New York.

Fuller MT (1998) Genetic control of cell proliferation and differentiation in *Drosophila* spermatogenesis. Semin. Cell Dev. Biol. 9: 433-444.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD and Bairoch A. Protein identification and analysis tools of the Expasy server, in The proteomics protocols handbook (Edited by Walker JM) pp571-607, Humana Press Inc, Totowa, NJ.

Gonczy P, Matunis E and DiNardo S (1997) *bag-of-marbles* and *benign gonial cell neoplasm* act in the germline to restrict proliferation during *Drosophila* spermatogenesis. Development 124: 4361-4371

Gould-Somero M, Holland L (1974) The timing of RNA synthesis for spermiogenesis in organ cultures of *Drosophila melanogaster* testes. Wilhelm Roux.Arch. 174: 133–148.

Greenspan RJ (1997) The basics of doing a cross, in Fly pushing, The theory and practice of *Drosophila* genetics pp 3-21. Cold Spring Harbor Press, New York.

Haerty W and Singh RS (2006) Gene regulation divergence is a major contributor to the evolution of Dobzhansky-Muller incompatibilities between species of *Drosophila*. Mol. Biol. Evol. 23: 1707-1714.

Haerty W, Jagadeeshan S, Kulathinal RJ, Wong A, Ravi Ram K, Sirot LK, Levesque L, Artieri CG, Wolfner MF, Civetta A, Singh RS (2007) Evolution in the fast lane: rapidly evolving sex-related genes in Drosophila. Genetics 177:1321-1335.

Haldane JBS (1922) Sex ratio and unisexual sterility in animal hybrids. J. Genet. 12: 101-109.

Hales KG and Fuller MT (1997) Developmentally regulated mitochondrial fusion mediated by a conserved, novel, predicted GTPase. Cell 90: 121-129.

Hecht NB (2000) Intracellular and intercellular transport of many germ cell mRNAs is mediated by the DNA- and RNA-Binding protein, Testis-Brain-RNA-Binding Protein (TB-RBP) Molecular Reproduction and Development 56: 252-253

Hiller M, Chen X, Pringle MJ, Suchorolski M, Sancak Y, Viswanathan S, Bolival B, Lin T-Y, Marino S and Fuller MT (2004) Testis-specific TAF homologs collaborate to control a tissue-specific transcription program. Development 131: 5297-5306

Hime GR, Brill JA, Fuller MT (1996) Assembly of ring canals in the male germline from structural components of the contractile ring. Journal of Cell Science 109: 2779-2788

Hollocher H and Wu C-I (1996) The genetics of reproductive isolation in the *Drosophila simulans* clade: X vs. autosomal effects and male vs. female effects. Genetics 143: 1243-1255.

Huber FK, Kaiser R, Sauter W, Schiestl FP (2005) Floral scent emission and pollinator attraction in two species of *Gymnadenia* (Orchidaceae). Oecologia 142: 564–75.

Hurst LD and Pomiankowski A (1991) Causes of sex ratio bias may account for unisexual sterility in hybrids: A new explanation of Haldane's rule and related phenomena. Genetics 128: 841-858.

Hwa JJ, Hiller MA, Fuller MT and Santel A (2002) Differential expression of the *Drosophila* mitofusin genes *fuzzy onions* (*fzo*) and *dmfn*. Mechanisms of Development 116: 213-216.

Jaenike J (2001) Sex chromosome meiotic drive. Annu. Rev. Ecol. Syst. 32: 25-49.

Jiggins CD, Naisbit RE, Coe RL, Mallet J. (2001) Reproductive isolation caused by colour pattern mimicry. Nature. 411: 302-305.

Kibbe WA (2007) OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res 35: W43-W46.

Kiger AA, Jones DL, Schulz C, Rogers MB, Fuller MT (2001) Stem cell self-renewal specified by JAK-STAT activation in response to a support cell cue. Science 294: 2542-2545.

Kulathinal R and Singh RS (1998) Cytological characterization of premeiotic versus postmeiotic defects producing hybrid male sterility among sibling species of the *Drosophila melanogaster* complex. Evolution 52: 1067-1079.

Lachaise D, David JR, Lemeunier F, Tsacas L and Ashburner M (1986) The reproductive relationships of *Drosophila sechellia* with *D. mauritiana*, *D. simulans*, and *D. melanogaster* from the Afrotropical region. Evolution 40: 262-271.

Lancefield DE (1929) A genetic study of crosses of two races or physiological species of *Drosophila obscura*. Zeitscrift ffir induktives Abstammungs- und Vererbungslehre, 52: 287-317.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.

Licht LE and Bogart JP (1985) Haldane's rule OK, but not always. Nature 317: 207.

Lin T-Y, Viswanathan S, Wood C, Wilson PG, Wolf N and Fuller MT (1996) Coordinate developmental control of the meiotic cell cycle and spermatid differentiation in *Drosophila* males. Development 122: 1331-1341

Lindsley DL and Tokuyasu KT (1980) Spermatogenesis, in The Genetics and Biology of *Drosophila* Vol.2 (Ashburner, M. and Wright, T. R. F., eds.) pp. 225–294. Academic, NewYork.

Lloyd JE (1966) Studies on the flash communication system in *Photinus* fireflies. Misc. Publ. Mus. Zool. Univ. Michigan. 130: 1-195.

Manly BJF (1991) Randomization and Monte Carlo methods in biology. Chapman and Hall, London.

McGregor JR, Xi R and Harrison DA (2002) JAK signaling is somatically required for follicle cell differentiation in *Drosophila*. Development 129: 705-717

McKearin DM and Spradling AC (1990) bag-of-marbles: a *Drosophila* gene required to initiate both male and female gametogenesis. Genes Dev. 4: 2242-2251

Michalak P and Noor MAF (2003) Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. Mol. Biol. Evol. 20:1070–1076.

Michalak P and Noor MAF (2004) Association of misexpression with sterility in hybrids of *Drosophila simulans* and *D. mauritiana.* J. Mol. Evol. 59:277-282.

Michalak P and Ma D (2008) The acylphosphatase (*Acyp*) alleles associate with male hybrid sterility in *Drosophila*. Gene 416: 61-65

Moehring AJ, Teeter KC, Noor MAF (2007) Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. II. Examination of multiple-species hybridizations, platforms, and life cycle stages. Mol. Biol. Evol. 24:137–145.

Montchamp-Moreau C and Joly D (1997) Abnormal spermiogenesis is associated with the X-linked sex-ratio trait in *Drosophila simulans*. Heredity 79:24-30.

Moore JA (1946) Incipient intraspecific isolating mechanisms in *Rana pipens*. Genetics 31: 304-326.

Morales CR, Wu XQ, Hecht NB (1998) The DNA/RNA-Binding Protein, TB-RBP, moves from the nucleus to the cytoplasm and through intercellular bridges in male germ cells. Development Biology 201: 113-123.

Mozdy AD and Shaw JM (2003) A fuzzy mitochondrial fusion apparatus comes into focus. Nature Reviews Molecular Cell Biology 4: 468-478.

Ohlstein B, Lavoie CA, O. Vef, E. Gateff, D. M. McKearin (2000) The *Drosophila* cystoblast differentiation factor, *benign gonial cell neoplasm*, is related to DExH-box proteins and interacts genetically with *bag-of-marbles*. Genetics 155: 1809-1819

Olivieri G and Olivieri A (1965) Autoradiographic study of nucleic acid synthesis during spermatogenesis in *Drosophila melanogaster*. Mutat. Res. 2:366–380.

Orr HA (1997) Haldane's rule. Annu. Rev. Ecol. Syst. 28:195-218.

Orr HA and Irving S (2005) Segregation distortion in hybrids between the Bogota and USA subspecies of *Drosophila pseudoobscura* Genetics169:671-682.

Palopoli MF and Wu C-I (1994) Genetics of hybrid male sterility between Drosophila sibling species: A complex web of epistasis is revealed in interspecific studies. Genetics 138: 329-341

Pavlidis P, Li Q, Noble WS (2003) The effect of replication on gene expression microarray experiments. Bioinformatics 19:1620-16207

Perez DE, Wu C-I, Johnson NA and Wu ML (1993). Genetics of reproductive isolation in the *Drosophila simulans* clade: DNA marker-assisted mapping and characterization of a hybrid-male sterility gene, Odysseus (Ods). Genetics 134:261-275.

Perez DE and Wu C-I (1995). Further characterization of the *Odysseus* locus of hybrid sterility in *Drosophila*: One gene is not enough. Genetics 140:201-206.

Phadnis N and Orr HA (2009) A single gene causes both male sterility and segregation distortion in *Drosophila* hybrids. Science 323:376-379.

Price CS, Kim CH, Gronlund CJ, Coyne JA (2001) Cryptic reproductive isolation in the *Drosophila simulans* species complex. Evolution 55:81-92

Ranz JM, Namgyal K, Gibson G, Hartl DL (2004) Anomalies in the expression profile of interspecific hybrids of *Drosophila melanogaster* and *Drosophila simulans*. Genome Res. 14:373-379

Sandler L and Novitski E (1957) Meiotic drive as an evolutionary force. Am. Nat. XCI: 105-110.

Sandler L, Hiraizumi Y, Sandler I (1959) Meiotic drive in natural populations of *Drosophila melanogaster*. I. The cytogenetic basis of segregation-distortion 44: 233-250.

Sawamura K (1996) Maternal effect as a cause of exceptions for Haldane's rule. Genetics 143: 609-611.

Silver DL and Montell DJ (2001) Paracrine signaling through the JAK/STAT pathway activates invasive behavior of ovarian epithelial cells in *Drosophila*. Cell 107: 831-841.

Singh RS (1990) Patterns of species divergence and genetic theories of speciation. In Population biology: Wohrmann K, Jain SK (eds). Springer-Verlag, Berlin, pp 231-265.

Singh RS, Kulathinal RJ (2005) Male sex drive and the masculinization of the genome. Bioessays 27:518-525.

Sturtevant AH (1920) Genetic studies on *Drosophila simulans*. I. Introduction. Hybrids with *Drosophila melanogaster*. Genetics 5: 488-500.

Tao Y, Chen S, Hartl DL and Laurie CC (2003a) Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. I. Differential accumulation of hybrid male sterility effects on the X and autosomes. Genetics 164: 1383-1397.

Tao Y, Zeng Z-B, Li J, Hartl DL and Laurie CC (2003b) Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. II. Mapping hybrid male sterility loci on the third chromosome. Genetics 164:1399-1418.

Tao Y and Hartl DL (2003) Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. III. Heterogeneous accumulation of hybrid incompatibilities, degree of dominance, and implications for Haldane's rule. Evolution 57:2580-2598.

Taylor MJ and Short RV (1973) Development of the germ cells in the ovary of the mule and hinny. J. Reprod. Fert. 32:441-445.

Thomas S and Singh RS (1992) A comprehensive study of genetic variation in natural populations of *Drosophila melanogaster*. VII. Varying rates of genic divergence as revealed by 2-dimensional electrophoresis. Mol. Biol. Evol. 9:507-525.

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

True JR, Weirt BS, Laurie CC (1996) A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans*. Genetics 142:819-837.

Ventela S, Toppari J, Parvinen M (2003) Intercellular organelle traffic through cytoplasmic bridges in early spermatids of the rat: Mechanisms of haploid gene product sharing. Molecular Biology of the Cell 14:2768-2780

Wakimoto BT, Lindsley DL, Herrera C (2004) Toward a comprehensive genetic analysis of male fertility in *Drosophila melanogaster*. Genetics 167:207-216.

White-Cooper H, Leroy D, MacQueen A, Fuller MT (2000) Transcription of meiotic cell cycle and terminal differentiation genes depends on a conserved chromatin associated protein, whose nuclear localisation is regulated. Development 127: 5463-5473

White-Cooper H, Doggett K, Ellis RE (2009) The evolution of spermatogenesis. In: *Sperm Biology: An Evolutionary Perspective*. Ed: TR Birkhead, DJ Hosken and S. Pitnick. Pp. 151-183

White-Cooper H (2010) Molecular mechanisms of gene regulation during *Drosophila* spermatogenesis. Reproduction 139: 11-21

Wong MD, Jin Z, Xie T (2005) Molecular mechanisms of germline stem cell regulation. Annual Review of Genetics. 39:173-195

Wu C-I and Beckenbach AT (1983) Evidence for extensive genetic differentiation between the sex-ratio and the standard arrangement of *Drosophila pseudoobscura* and *D. persimilis* and identification of hybrid sterility factors. Genetics 105:71-86.

Wu C-I and Davis AW (1993) Evolution of postmating reproductive isolation: The composite nature of Haldane's rule and its genetic bases Am. Nat 142:187-212.

Wu C-I, Johnson NA, Palopoli MF (1996) Haldane's rule and its legacy: why are there so many sterile males? Trends Ecol. Evol. 11:281-284.

You FM, Huo N, Gu YQ, Luo MC, Ma Y, Hane D, Lazo GR, Dvorak J, Anderson OD (2008) BatchPrimer3: a high throughput web application for PCR and sequencing primer design. BMC Bioinformatics 9:253.

Zeng L-W and Singh RS (1993) A combined classical genetic and high resolution two-dimensional electrophoretic approach to the assessment of the number of genes

affecting hybrid male sterility in *Drosophila simulans* and *Drosophila sechellia* Genetics 135:135-147.

Appendix I – Standard cornmeal-molasses-yeast-agar (CMYA) medium.

Ingredient	**Quantity**
Cornmeal	65 g
Brewers Yeast (debittered)	13 g
Agar	6.5 g
Cold Water	170 ml
Water to boil	760 ml
Molasses	45.5 ml
10% Tegosept*	20 ml
Concentrated Propionic acid	5 ml

* - 50 g methyl 4-hydroxybenzoate in 500 ml of 95% denatured ethanol.

Protocol:

Pour water to boil into a steel vessel and bring to boil on a hot plate. Mix cornmeal, yeast, agar with cold water to make into slurry. Breaks any lumps formed and pour slurry to boiling water. Stir constantly to prevent burning until the mixture come to second boil, then remove it from heat (Debittered yeast is inactive and hence no leavening). Add molasses and cool to 60-65ºC, add Tegosept and propionic acid. Mix well and dispense into autoclaved vials and bottles with peristaltic pump, cover with cheese cloth until it solidifies. Yields about one hundred 28.5mm × 95mm vials or eighteen 237 ml bottles.

Appendix II – 1X Phosphate-Buffered Saline (PBS) solution.

Ingredient	**Quantity**
Sodium chloride (NaCl)	8 g
Potassium chloride (KCl)	0.2 g
Sodium phosphate (Na_2HPO_4)	1.44 g
Potassium phosphate (KH_2PO_4)	0.24 g

Dissolve all salts in 800 ml DDH_2O. Adjust pH to 7.2 and make up to 1 liter with DDH_2O. Sterilize by autoclaving and store at room temperature

Appendix III – 10X TBE Buffer.

Ingredient	Quantity
Tris	108 g
Boric acid	55 g
EDTA	5.85 g

Dissolve all salts in 800 ml of DDH_2O and make it up to a final volume of 1L. Sterilize by autoclaving and store at room temperature. Dilute 100 ml of 10X TBE with 900 ml of DDH_2O to make 1X TBE buffer.

Appendix IV – Protein Sample (PS) buffer.

Ingredient	**Quantity**
DTT	0.1 g
CHAPS	0.4 g
Urea	5.4 g
Ampholyte	500 µl
DDH_2O	6 ml

The final volume of the buffer will be approximately 6.5 ml. Shake well and divide into several batches of 100 µl and store at -70º C until further use.

Appendix V – Separating Gel.

To cast a separating gel, prepare separating buffer by dissolving 90.8 g of Tris in 400 ml of DDH_2O and adjust the pH to 8.8 using 3M or 1M HCl. Bring the final volume of the buffer to 500 ml by adding DDH_2O, sterilize by autoclaving and store at room temperature. Once after making a stock of separating buffer, cast the separating gel using the following ingredients.

Ingredient	**Quantity**
30% / 0.8% Acrylamide / Bis acrylamide	8.3 ml
DDH_2O	6.5 ml
Separating Gel Buffer	4.95 ml
10% SDS	198 µl
TEMED	20 µl
40% APS	145.3 µl

Add all ingredients in a sequential order except 40% APS and mix thoroughly. Add required amount of 40% APS exactly prior pouring the gel into the casting frame, because TEMED and APS cause the buffer to polymerize rapidly.

Appendix VI – Stacking Gel.

To cast a separating gel, prepare stacking buffer by dissolving 90.8 g of Tris in 400 ml of DDH_2O and adjust the pH to 8.8 using 3M or 1M HCl. Bring the final volume of the buffer to 500 ml by adding DDH_2O, sterilize by autoclaving and store at room temperature. Once after making a stock of stacking buffer, cast the stacking gel using the following ingredients.

Ingredient	**Quantity**
30% / 0.8% Acrylamide / Bis acrylamide	850 µl
DDH_2O	2.7 ml
Stacking Gel Buffer	1.25 ml
10% SDS	50 µl
TEMED	8 µl
40% APS	15.6 µl

Add all ingredients in a sequential order except 40% APS and mix thoroughly. Add required amount of 40% APS exactly prior pouring the gel into the casting frame, because TEMED and APS cause the buffer to polymerize rapidly.

Appendix VII – Tank Buffer (10X).

Ingredient	Quantity
Tris	30 g
Glycine	144 g
10% SDS	100 ml

Dissolve all ingredients in 800 ml of DDH_2O and adjust the pH to 8.3 by adding 3M or 1M HCl. Make up to a final volume of 1L and do not autoclave. Dilute 100 ml of this 10X tank buffer with 900 ml of DDH_2O to make 1L of 1X tank buffer solution.

Appendix VIII – 2X Sample Buffer.

Ingredient	**Quantity**
Stacking Gel Buffer	2.5 ml
Glycerol	2 ml
Sodium Dodecyl Sulphate (SDS)	400 mg
Bromophenol Blue (BPB)	100 µg (or few grains)

Dissolve all salts are dissolved in 8 ml of DDH_2O and make it up to a final volume of 10 ml. Shake well and divide into batches of 1 ml and store at -70º C until further use.

Appendix IX – Staining SDS gel using Coomassie blue.

Fixing solution:

Dissolve 400 ml of 95% ethanol and 100 ml of glacial acetic acid in 500 ml of DDH_2O. Store solution at room temperature.

Destaining solution:

Dissolve 250 ml of ethanol and 80 ml of glacial acetic acid in 670 ml of DDH_2O. Store solution at room temperature.

Staining solution:

Dissolve 1 tablet of PhastGel Coomassie Blue R in 400 ml of destaining solution in a boiling water bath. Filter the solution to remove undissolved crystals and store at room temperature.

Protocol:

Once the electrophoresis is complete, gently remove the gel from the casting frame and soak in fixing solution for 30 min in a clean glass tray. Discard the fixing solution and rinse the gel with DDH_2O. Soak the gel in the filtered staining solution for 30 min. After incubation discard the staining solution and rinse the gel twice with DDH_2O. Soak the stained gel in destaining solution overnight or until the gel is clear displaying protein bands. Fresh destaining solution can be replaced once the solution gets colored. All incubations must be performed on a clean glass tray over a platform shaker at low speed.

Appendix X – Equilibration stock solution.

Ingredient	Quantity
Tris – HCl (pH 6.8)	20 ml
Urea	72 g
Glycerol	60 ml
SDS	2 g

Protocol:

Dissolve all salts are dissolved in 150 ml of DDH_2O, make up to a final volume of 200 ml using DDH_2O and store at room temperature. Do not autoclave the stock solution. Prior equilibrating IPG strips prepare equilibration buffer no. 1 and no.2 freshly as indicated below. These buffers cannot be stored for further use.

A. Equilibration buffer no.1: Add 25 mg of DTT to 10 ml of equilibration stock solution and vortex to dissolve completely.

B. Equilibration buffer no.2: Add 0.45 g of iodoacetamide and few grains of bromophenol blue to 10 ml of equilibration stock solution and vortex to dissolve completely.

Appendix XI – Electrode buffers.

A. Anode Buffer:

Ingredient	Quantity
Tris	13.63 g
SDS	1 g

Dissolve Tris and SDS in 200 ml of DDH_2O and adjust to pH 6.6 with glacial acetic acid. Make up to the final volume of 250 ml using DDH_2O and do not autoclave.

B. Cathode Buffer:

Ingredient	Quantity
Tris	9.7 g
Tricine	14.3 g
SDS	0.6 g

Dissolve all ingredients in 80 ml of DDH_2O and adjust to pH 7.1 with glacial acetic acid. Make up to the final volume of 100 ml using DDH_2O and do not autoclave.

Appendix XII – Alignments of *D. simulans*, *D. mauritiana* spermatogenesis candidate genes and two reference genes. The *D. simulans* sequence (sim) was obtained from the Flybase database and the *D. mauritiana* (mau) sequence was generated through my sequencing data. The locations of the primers used for qRT-PCR are underlined. Introns flanked by *bam* and *RpL32* primers are shown in lower case font.

hop

```
sim TGAGTGTGGAGCGTTTGAAGTGGCACTATGTGCATCAGGTCTCCCACCTGGCGCCCACCTATATGAC
mau TGAGTGTGGAGCGTTTGAAGTGGCACTATGTGCATCAGGTCTCCCACCTGGCGCCCACCTATATGAC
    ************************************************************

sim CGAACAGTTTACCTGCACCGTTCAGTATCTGCCCAACGAGGAGGTGGCCCGCGGCAGCGGATCCATC
mau CGAACAGTTTACCTGCACCGTTCAGTATCTGCCCAACGAGGAGGTGGCCCGCGGCAGCGGATCCATC
    ************************************************************

sim GGCACCAGTCTGGCCCACTCGACGTCGTCGCTGTCCAGTTCCGGTTCCACCAACACTCTGTCCA
mau GGCACCAGTCCGGCCCACTCGACGTCGTCGCTGTCCAGTTCCGGTCCCACCAACACTCTGTCCA
    ********** ************************************************
```

piwi

```
sim ATTGCGAAGAGCACACGAGATCGCAAGAGGGCCTACGGAGCATTGATTGCCTCAATGGATCTACAGC
mau ATTGCGAAGAGCACACGAGATCGCAAGAGGGCCTACGGAGCATTGATTGCCTCAATGGATCTACAGC
    ************************************************************

sim AAAACTCCACGTACTTCAGCACGGTCACGGAGTGCAGTGCCTTTGATGTGCTCGCAAACACCCTTTG
mau AAAACTCCACGTACTTCAGCACGGTCACGGAGTGCAGTGCCTTTGATGTGCTCGCAAACACCCTTTG
    ************************************************************

sim GCCTATGATAGCAAAGGCCCTGCGCCAGTATCAACTAGAGCATAAGAAGCTGCCATCTCGAATCGTA
mau GCCTATGATAGCAAAGGCCTTGCGCCAATATCAACTAGAGCATAAGAAGCTGCCATCTCGAATCGTA
    ******************* ******* ***************************************

sim TTTTATCGGGACGGG
mau TTTTATCGGGACGGG
    ***************
```

bam

```
sim CGCAATCGAAACGGAAACTCGGGGATCAATGCGGACAAGTgtaagctgtagattttcaagcaaccat
mau CGCAATCGAAACGGAAACTCGGGGAGCAATGCGGACAAGTgtaagctgtagattttcaagcaaccat
    ************************* ************************************

sim tcagttattcctgcaacgattttattcattatagTCCATGCTCAGCTCATGGAGAGATTGCTGATTG
mau tcagttattcctgcaacgattttattctttacagTCCATGCTCAGCGCATGGAGAGATTGCTGATTG
    *************************** *** ************** ********************

sim GTCTGCGCGATTGGATCAAGGCTGCGCATCTCAGTGTGCACGTGTTTAACTGGGAAATGGATCTGGA
mau GTCTGCGCGATTGGATCAAGGCTGCGCATCTCAGTGTGCACGTGTTTAACTGGGAAATGGATCTGGA
    ************************************************************
```

```
sim GCACCGCTACTCGGGGGCCATGACCGAAAGCCACAAGTCGTTGACCGAGCGGGCGATCCTTTTGTCT
mau GCACCGCTACTCGGGGGCCATGACCGAAAGCCACAAGTCGTTGACCGAGCGGGCGATCCTTTTGTCT
    ************************************************************************

sim GGTGCCG
mau GGTGCCG
    *******
```

bgcn

```
sim ACGGTGGCAATAACGGAACTTGCATTCGACGCGCGCAGTGCGGCGCACTTCGCTTGATTTTCTGTTC
mau ACGGTGGCAATAACGGAACTTGCATTCGACGCGCGCAGTGCGGCGCACTTCGCTTGATTTTCTGTTC
    ************************************************************************

sim TGTTTCTTCTGTTTGAATCTTCGCCGAAATGAACCACATCATCCAGGACAAGTATATTCCGCAGCAG
mau TGTTTCTTCTGTTTGAATCTTCGCCGAAATGAACCACATCATTCAGGACAAGTACATTCCGCAGCAG
    ****************************************** *********** ************

sim CTGCTCTACTTCTTGGCGGGCCGGCGCTGCTGCCAGCAGTTCCCTTGCACATTCCG
mau CTGCTCTACTTCTTGGCGGGCCGGCGCTGCTGCCAGCAGTTCCCTTGCACATTCCG
    ************************************************************
```

can

```
sim TTCGCTTGTGGTGCCTTCTGTCCTGGAGTTGTGTGGTTATTTTTCCCGGCCATTTGGCACCGGTTTG
mau TTCGCTTGTGGTGCCTTCTGTCCTGGAGTTGTGTGGTTATTTTTCCCGGCCATTTGGCACCGGTTTG
    ************************************************************************

sim CTTTGTTGTTTTTGCACCAAGGGGGTATTATTTCGCCACCGCATCGGATGATTGTACGGCGAGAG
mau CTTTGTTGTTTTTGCACCAAGGGGGTATTATTTCGCCACCGCATCGGATGATTGTACGGCGAGAG
    ************************************************************************
```

sa

```
sim AAAGCACCGGAGACACAAGACCAGGACTTTCACAGTGGACCACCACCCATACTGAGCTCCACAAAGG
mau AAAGCACCGGAGACACAAGACCAGGACTTTCCACAGTGGGCCACCACCCATGCTGAGCTCCACAAAGG
    ***************************** ******************** ****************

sim CCATGGTATTGGCCTCCACCGCGTACATTCCGGACTATTTGCCACCATTTCCAGG
mau CCATGGTATTGGCCTCCACCACGTACATTCCGGACTATTTGCCACCATTTCCAGG
    ******************** ***********************************
```

fzo

```
sim AAGCTCTCGCGTCCAAATCTCTTTATACTCAACAATCGATGGGATAAGGCCAGCAGTATGGAGCCGG
mau AAGCTCTCGCGTCCAAATCTCTTTATACTCAACAATCGATGGGATAAGGCCAGCAGTATGGAGCCGG
    ************************************************************************

sim AAATGGAGCAGAAGGTAAAGGATCAGCATATGGAACGCTGCGTTAATCTGCTAGTGGATGAGTTAGG
mau AAATGGAGCAGAAGGTAAAGGATCAGCATATGGAACGCTGCGTTAATCTGCTAGTGGATGAGTTAGG
    ************************************************************************

sim TGTTTATTCCACGGCG
mau TGTTTATTCCACGGCG
    ****************
```

Bruce

```
sim TTGCCGGAACTTGGATAGGCTATGCTGTTGCTCTCGACAGGGCGGCCTGCTCTTCTACTCGCTAAGC
mau TTGCCGGAACTTGGATAGGCTATGCTGTTGCTCTCGACAGGGCGGCCTGCTCTTCTACTCGCTAAGC
    ************************************************************************
```

```
sim GAGGGGGAGAACGATTCTGGAGACGAACTTCTGGAAATGGACGATGACTGCAGTACCACATTAACAC
mau GAGGGGGAGAACGATTCTGGAGACGAACTTCTGGAAATGGACGATGACTGCAGTACCACATTAACAC
    ************************************************************************

sim AGGCAGCGGA
mau AGGCAGCGGA
    **********
```

Act5C

```
sim CCGTGAGAAGATGACCCAGATCATGTTCGAGACCTTCAACACACCCGCCATGTATGTGGCCATCCAG
mau CCGTGAGAAGATGACCCAGATCATGTTCGAGACCTTCAACMCWCCCGCCATGTATGTGGCCATCCAG
    ****************************************** * ************************

sim GCTGTGCTCTCGCTGTACGCCTCCGGTCGTACCACCGGTATCGTTCTGGACTCCGGCGATGGTGTCT
mau GCYGTGCTCTCGCTGTACGCCTCCGGTCGTACCACCGGTATYGTTCTGGACTCCGGRGATGGTGTCT
    ** ************************************ ************** **********

sim CCCACACCGTGCCCATCTACGAGGGTTATGCCCTTCCCCATGCCATCCTGCGTCTGGATCTGGCTGG
mau CCCACACCGTGCCCATCTATGAGGGTTATGCCCTGCCCCAYGCCATCCTGCGTCTGGATCTGGCTGG
    ******************* ************** ***** *************************

sim TCGCGATTTGACCGACTACCTGATGAAGATCCTGACCG
mau TCGCGATTTGACCGACTACCTGATGAAGATCCTGACCG
    *************************************
```

RpL32

```
sim TACAGGCCCAAGATCGTGAAGAAGCGCACCAAGCACTTCATCCGCCACCAGTCGGATCGATATGCTA
mau TACAGGCCCAAGATCGTGAAGAAGCGCACCAAGCACTTCATCCGCCACCAGTCGGATCGATATGCTA
    ************************************************************************

sim AGCTGTCGgtgagtgcctacgacgattgtgccaaaagcccgtgtttaatccacatgtctccttgcag
mau AGCTGTCGgtgagtgcctacgaggattgtgccaaaagcccgtgtttaatccacatgtctccttgcag
    ********************** ****************************************

sim CACAAATGGCGCAAGCCCAAGGGTATCGACAACAGAGTGCGTCGCCGCTTCAAGGGACAGTATCTGA
mau CACAAATGGCGCAAGCCCAAGGGTATCGACAACAGAGTGCGTCGCCGCTTCAAGGGACAGTATCTGA
    ************************************************************************

sim TGCCCAACATCGGTTACGGATCGAACAAGCGCACCCGCCACATGCTGCCCACCGGATTCAAGAAGTT
mau TGCCCAACATCGGTTACGGATCGAACAAGCGCACCCGCCACATGCTGCCCACCGGATTCAAGAAGTT
    ************************************************************************

sim CCTGGTGCACAACGTGCGCGAGCTGGAGGTCCTGCTCATGCAGAACCGCGTCTACTGCGGCGAGATC
mau CCTGGTGCACAACGTGCGCGAGCTGGAGGTCCTGCTCATGCAGAACCGCGTCTACTGCGGCGAGATC
    ************************************************************************

sim GCCCACGGCGTCTCTTCCAAGAAGCGCAAGGAGATCGTCGAGCGCGCCAAGCAGCTGTCGGTCCGCC
mau GCCCACGGCGTCTCTTCCAAGAAGCGCAAGGAGATCGTCGAGCGCGCCAAGCAGCTGTCGGTCCSCC
    ************************************************************** **

sim CGGTCCGCCTCACCAACCCCAACGGT
mau CGGTCCSCCTCACCAACCCCAACGGT
    ****** ********************
```