EVOLUTIONARY AND FUNCTIONAL ANALYSIS OF THE DROSPHILA *BAG OF MARBLES* GENE

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Heather Ann Flores

January 2013

EVOLUTIONARY AND FUNCTIONAL ANALYSIS OF THE DROSPHILA *BAG OF MARBLES* GENE

Heather Ann Flores, Ph. D.

Cornell University 2013

Reproduction is fundamentally important for organismal fitness. While it is critical that each step in reproduction proceed correctly, many of the genes that regulate these processes are adaptively evolving in response to both internal and external selective pressures. We have previously shown that the *bag of marbles* (*bam*) and *benign gonial cell neoplasm* (*bgcn*) genes are evolving under rapid, adaptive evolution in *D. melanogaster* and *D. simulans*. These genes act at the earliest stages of reproduction, germline stem cell (GSC) regulation.

In the first study, I expand population genetic analyses to genes that play a role in GSC regulation and/or genetically interact with *bam*. I find that multiple GSC regulatory genes show evidence of rapid, adaptive evolution in *D. melanogaster* and/or *D. simulans*. These genes play very different roles in the regulation of GSCs, and their diverse functions and expression patterns suggest that multiple selective pressures are acting to drive the evolution of GSC regulatory genes.

In a second study, I focused on *bam* to determine the functional consequences of *bam*'s sequence divergence. I developed a transgenesis assay to determine the ability of a *bam* orthologs from *D. melanogaster* and *D. simulans* to rescue the male and female sterility of a *D. melanogaster bam* mutation. I found that while *D. simulans* can rescue *D. melanogaster bam* male sterility, it fails to fully rescue female

sterility suggesting that the selective pressure driving the evolution of *bam* is present in the female germline.

In the final study, I investigated the hypothesis that the bacterial endosymbiont, *Wolbachia pipientis*, is driving the adaptive evolution of *bam*. I found that *Wolbachia* and *bam* genetically interact by showing that *Wolbachia* can rescue the fertility defects of a *bam* hypomorph. I found that *D. melanogaster*-specific *Wolbachia* is unable to accumulate as well in the presence of a *D. simulans bam* ortholog which is consistent with a model that *Wolbachia* is co-evolving with *bam* and supports the hypothesis that *Wolbachia* is driving the adaptive evolution of *bam*.

BIOGRAPHICAL SKETCH

Heather Flores was born on March 17th, 1982 in Scottsbluff, Nebraska. Heather graduated from Scottsbluff High School in 2000 as a class valedictorian. In the fall of 2000, she began classes at the University of Nebraska-Lincoln where she majored in biology and mathematics and was part of the University Honors Program. Heather started research in applied mathematics. However, after taking genetics courses and participating in genetics research, she realized her passion was in the field of genetics. Heather graduated in May 2005. In fall of 2005 she entered the graduate field of Genetics & Development at Cornell University and joined the labs of Drs. Dan Barbash and Charles Aquadro. Upon completion of her dissertation, Heather continued on to a postdoctoral position in the lab of Dr. Scott O'Neill at Monash University in Melbourne, Australia.

To A.H.W.

ACKNOWLEDGMENTS

I owe my two advisors, Dan Barbash and Chip Aquadro much gratitude. The two of them allowed me to start a project that was new to both labs and have allowed me to have an immense amount of independence and input about where to take the project. They were also incredibly supportive of allowing me to attend meetings or receive training in labs whose expertise was in the Drosophila germline to gain the necessary expertise to work on my dissertation project. My committee members, Mariana Wolfner and Andy Clark provided invaluable insight on my project. Mariana, in particular, has been a great sounding board for questions both related and unrelated to my dissertation and provided in-depth comments on this dissertation which ultimately made it better.

This project would not have been possible if not for the advice and reagents of many who study the Drosophila germline. Dennis McKearin, Jean Maines, and Margaret Fuller provided endless advice in addition to reagents, many of which were unpublished. Ting Xie welcomed me into his lab at the Stower's Institute to learn relevant techniques, and his student Lei Pan was an excellent teacher.

My graduate work would not have been nearly as fulfilling without the camaraderie of members from both labs. I was fortunate to follow Shamoni Maheshwari and Nick Brideau in the Barbash lab, allowing me to use many of the techniques they had already optimized in this dissertation. Additionally, both have been great friends providing many stimulating conversations about science, college football, cats, and of course, gossip. Patrick Ferree provided advice on and shared his love and interest in *Wolbachia* with me. I can only hope one day to be as intelligent, driven, and organized as Nadia Singh. Through our friendship, I ventured into the areas of cooking classes, swimming, sewing, cake decorating and of course, the not-

so-successful foray into pottery. Vanessa Bauer DuMont has been an incredible friend and lab mate. She constantly entertains my questions and provides valuable insight and advice into my project. Her positive attitude has been very helpful these last few months with the stress of graduate school. There are few words to describe Satyaki Rajavasireddy. I cannot count the many hours we have spent arguing about experiments, politics, and crazy science ideas. He has been a constant sounding-board for experiments and provides endless entertainment in lab with his singing and bad jokes.

My time in Ithaca would not be as enjoyable without the many friends I have made here. I will always miss football Saturdays, cooking, and gossiping with SaraH and Pat Zanders and Maho Shibata. I am especially indebted to SaraH for providing assistance with Southern Blots and constantly keeping me on my toes when it comes to science.

Some of my biggest appreciation goes to Kyle Molina who has stood by me through all the science, providing unconditional love and support. He moved across the country and away from family to help support my graduate school endeavor. He has taken care of me and reminds me to take care of myself. He has probably spent more time at Biotech than many graduate students.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

A central goal in evolutionary biology is to identify the role that adaptation has played in creating the natural variation present both within and between species. There are two general approaches to studying adaptation. In a phenotype-first approach (also known as top-down approach), one identifies a phenotype that likely has an adaptive basis and moves toward identifying the genes that underlie that particular phenotype. This approach can be advantageous if the genetic architecture underlying the phenotype of interest is well-known. This approach has been very successful for identifying not only the genes but in some cases the particular polymorphisms; for example, genes underlying adaptive coat color in different mouse species (Nachman et al., 2003; Hoekstra et al., 2006). A drawback of this approach is that it is biased to identifying phenotypes that have an obvious adaptive basis. In a genotype-first approach (also known as bottom-up approach), one looks for footprints of adaptation within a genome and moves to identifying the affected phenotype. This approach has the advantage of being unbiased in that one simply allows the genome to direct oneself to genes that have experienced adaptation. This approach is commonly used in many taxa to identify adaptively evolving genes (Bustamante et al., 2005; Begun et al., 2007; Drosophila 12 Genomes Consortium et al., 2007; Kosiol et al., 2008; Larracuente et al., 2008). The main drawback of this approach is that it may be difficult to identify the affected phenotype if the adaptively evolving genes have no known functions. Additionally, the particular system in which adaptively evolving genes are involved may not have the proper tools needed to experimentally identify the affected phenotype.

A phenotype-first approach was used to discover that male accessory gland proteins are diverging under adaptive evolution. In Drosophila, males transfer proteins to females that influence her reproductive biology, a process that could lead to conflict between the reproductive strategies of males and females. It was hypothesized that male accessory gland proteins are the basis for these rapidly changing traits, and a phenotype-first approach was used to show that many of these genes are also rapidly evolving (Aguade et al., 1992; Tsaur et al., 1998; Swanson et al., 2001; Swanson et al., 2001). Soon after, genotype-first approaches were used to expand these analyses to a broader set of reproductive proteins in both males and females. These unbiased approaches discovered that many genes expressed throughout reproductive development in both males and females have diverged under adaptive evolution (Civetta and Singh, 1995; Good and Nachman, 2005; Civetta et al., 2006; Panhuis and Swanson, 2006; Bauer DuMont et al., 2007; Haerty et al., 2007; Kelleher et al., 2007; Li et al., 2009; Obbard et al., 2009; Kolaczkowski et al., 2011; Wong et al., 2012). Now, a focus of the field is to identify the adaptive phenotypes to which these genes contribute and to determine the selective pressures that drive the evolution of these genes.

The focus of my dissertation has been to identify the functional consequences of adaptive evolution of genes expressed at an early stage of reproduction, germline stem cell (GSC) regulation. We (Bauer DuMont et al., 2007) and others (Civetta et al., 2006) have shown that two of these genes, *bag of marbles* (*bam*) and *benign gonial cell neoplasm* (*bgcn*) are diverging under rapid, adaptive evolution in *D. melanogaster* and *D. simulans*. *bam* and *bgcn* are required for GSCs to differentiate

in the Drosophila ovary. The Drosophila ovary is comprised of units called ovarioles that are an assembly line for the developing oocyte. The GSCs are present at the most proximal end of each ovariole in a region called the germarium. GSCs reside in a niche environment made of both somatic and escort stem cells that are required to maintain their stem cell state (Lin, 1998; Decotto and Spradling, 2005; Kirilly and Xie, 2007). If the GSCs move even one cell-width away from the niche, they will differentiate. GSC daughter cells that do move away from the niche differentiate and undergo four mitotic divisions in synchrony to give rise to a 16-cell, interconnected cyst. One cell will become the oocyte, while the other 15 cells, termed nurse-cells, will provide necessary RNA and protein products to the oocyte. In spermatogenesis, the GSC daughter also undergoes four rounds of mitotic division; however, all 16 cells enter meiosis (Wong et al., 2005).

The germline stem cell regulatory system in Drosophila provides a valuable system in which to address questions of functional divergence. The regulation of germline stem cells is an incredibly active field of study, and many of the key regulators are known (Kirilly and Xie, 2007). This system has a plethora of genetic tools in addition to cytological reagents, making it easy to monitor oogenesis and spermatogenesis progression (Wong et al., 2005; Kirilly and Xie, 2007).

bam and *bgcn* are expressed at early stages of reproduction, and prior to work by Civetta *et. al*. (2006) and us (Bauer DuMont et al. 2007), no other strictly-early germline genes had been shown to evolve under adaptive evolution. It was not clear if *bam* and *bgcn* are unique in their adaptive evolution or if they are just two of many early reproductive genes adaptively evolving. The first approach I took to address

these questions was to perform a population genetic survey of other genes that either interact with *bam* or are expressed early in GSC development. If either of these classes of genes also showed adaptive evolution, these data might help us identify the function of *bam* that is under selection. The population survey described in Chapter 2 provided evidence that *bam* and *bgcn* are not the only GSC gene that are adaptively evolving. Half the genes we surveyed showed evidence of adaptive evolution in *D. melanogaster* and/or *D. simulans*. Our hope was that the functions of other adaptively evolving genes would help to identify a selective pressure acting on *bam* and *bgcn*. However, the genes we identified play very different roles in GSC regulation and are expressed in different tissues (i.e. germline and somatic tissue). In fact, their diverse functions suggest that multiple selective pressures are likely acting in the early germline to drive the adaptive evolution of GSC regulatory genes.

While the population genetics approach is important in establishing that many GSC regulatory genes are adaptively evolving, it was clear that it was limited in what more it could contribute to our understanding of the relevant selective pressures. Therefore, I decided to combine the population genetics analysis with a functional analysis, specifically focusing on the *bam* gene. In Chapter 3, I describe the generation of transgenes to assay the ability of *bam* orthologs from *D. melanogaster* or *D. simulans* to rescue the male and female sterility of a *D. melanogaster bam* null mutation. *bam* has multiple functions in early germline development, and our expectation was that the adaptive evolution of *bam* had driven the divergence of a subset of these functions. Once we identified the non-rescuing functions, we could utilize this information to make informed hypotheses about the selective pressures

acting on *bam* and potentially other genes involved in GSC regulation. This approach was quite successful in that there was a clear dichotomy in the ability of the transgenes to rescue specific phenotypes: the *D. simulans bam* ortholog could rescue male sterility but failed to fully rescue female sterility. This single observation was crucial in helping us to refine and begin to test specific hypotheses of selective pressures acting on *bam* in the female germline.

There are a number of established selective pressures acting on the female germline, many of which act downstream of where *bam* is expressed or involve functions for which *bam* is not known to participate. In Chapter 4, I test whether the maternally-inherited bacterial endosymbiont, *Wolbachia pipientis*, could be a selective pressure driving the evolution of *bam*. I find that *bam* and *Wolbachia* genetically interact, providing strong support for *Wolbachia* being an important selective pressure. While I would like to show direct evidence that *Wolbachia* is driving the adaptive evolution of *bam*, showing direct cause for historical events is incredibly difficult. Instead, I investigate two different models of how *Wolbachia*'s interaction with *bam* would drive it to adaptively evolve and find that the data are more consistent with the predictions of a model of antagonistic coevolution between *bam* and *Wolbachia*. The use of this integrated approach has not only broadened our view on the types of reproductive proteins that are adaptively evolving, but it has also broadened our view on the types of selective pressures acting throughout the germline that could drive the adaptation of reproductive genes.

CHAPTER 2

PERVASIVE ADAPTIVE EVOLUTION OF GENES INVOLVED IN THE REGULATION OF DROSOPHILA GERMLINE STEM CELLS[1]

2.1 Introduction

Genome-wide surveys of polymorphism and divergence are identifying surprisingly high levels of adaptive evolution among protein-coding genes (Bustamante et al., 2005; Begun et al., 2007; Drosophila 12 Genomes Consortium et al., 2007; Haerty et al., 2007; Macpherson et al., 2007; Larracuente et al., 2008). These discoveries raise two fundamental questions: what are the functional consequences for the genes under positive selection, and what are the selective pressures driving these changes? Reproduction and fertility are among the most important traits for organismal fitness. Many models and theoretical studies have proposed that germline and fertility-related genes will be targeted for selection by various processes, and empirical evidence has documented rapid evolution and in some cases positive selection in numerous genes known or proposed to be involved in male fertility (Tsaur et al., 1998; Begun et al., 2000; Swanson et al., 2001; Swanson et al., 2004; Clark and Swanson, 2005; Haerty et al., 2007), female reproductive tract function (Lawniczak and Begun, 2007; Prokupek et al., 2008; Kelleher and Markow, 2009), host defense against segregation distorters (Presgraves, 2007; Phadnis and Orr,

[1] This work was done in collaboration with Vanessa L. Bauer DuMont, Aalya Fatoo, Diana Hubbard, and Mohammed Hijji. Author contributions are as follows: H.A.F., V.L.B.D., A.F., D.H., and M.H, performed population sequencing. H.A.F. and V.L.B.D. analyzed the data and wrote the paper.

2009), and sperm-egg interactions (Swanson and Vacquier, 1995; Swanson et al., 2001; Aagaard et al., 2010). Most of these genes are expressed at the latter stages of gametogenesis and are often associated with meiosis or interactions between gametes.

Recently, Civetta *et al*. (2006) and we (Bauer DuMont et al., 2007) independently discovered that genes expressed in earlier stages of gametogenesis, specifically germline stem cell (GSC) regulation, also show evidence of adaptive evolution. One of these genes, the *bag of marbles* (*bam*) gene, is under strong positive selection with an astonishing 59 nonsynonymous substitutions among 442 codons between two closely related fruit fly species, *Drosophila melanogaster* and *D. simulans* (Civetta et al., 2006; Bauer DuMont et al., 2007). Given that the gene *benign gonial cell neoplasm* (*bgcn*) acts together with *bam* in GSC differentiation, we examined variation at *bgcn* and found that it is also evolving under positive selection in these two species (Bauer DuMont et al., 2007).

Germline stem cells (GSCs) produce the cells that will further develop to form either eggs or sperm. This production occurs throughout an animal's life, and thus, GSCs must be carefully regulated. GSCs are maintained in a microenvironment called the stem cell niche that is located in the proximal end of the Drosophila ovary or the apical end of the testis. GSCs must remain in the niche to retain a stem cell state (Fuller and Spradling, 2007; Spradling et al., 2011). When a GSC undergoes an asymmetric division, one of the daughter cells moves out of the niche and differentiates (Figure 2.1). This differentiated daughter cell undergoes four synchronized, mitotic divisions with incomplete cytokinesis to give rise to a 16 cell-cyst. In females, one of the 16 cells will become the future oocyte whereas the other

15 become nurse cells that provide mRNAs and proteins for the developing oocyte. In males, all 16 cells enter meiosis giving rise to 64 spermatids (reviewed in Wong et al., 2005; Kirilly and Xie, 2007).

The gene circuitry involved in GSC regulation in females is described below briefly for functional context and visualized in Figure 2.1. It should be noted, however, that many of these genes function somewhat differently in the male GSCs (e.g. Gonczy et al., 1997; Gilboa and Lehmann, 2004; Kawase et al., 2004; Song et al., 2004; Fuller and Spradling, 2007; Insco et al., 2009). GSC regulation is carefully controlled by both intrinsic signals within the GSC and extrinsic signals from the niche. To receive extrinsic signals, GSCs remain physically attached to the niche through adherens junctions (Song and Xie, 2002). The gap junction protein Zero population growth (Zpg) is present in the cytoplasmic membrane of both GSCs and niche cells and is required for the maintenance of GSCs through the sharing of small molecules and signals between the niche and GSC (Tazuke et al., 2002; Gilboa et al., 2003). A single gene, *bag of marbles* (*bam*), is a switch to allow for GSC differentiation, and therefore has its expression repressed in the GSCs (McKearin and Ohlstein, 1995). The niche provides extrinsic signals that repress the transcription of *bam* (Song et al., 2004). However, this signaling is quickly dissipated and thus will only repress *bam* in those cells that are in physical contact with the cap cells (Wong et al., 2005; Xia et al., 2010). Multiple, partially redundant signaling mechanisms repress *bam* expression in the GSC. The repression is predominantly due to the bone morphogenic protein (BMP) signaling pathway (Xie and Spradling, 1998; Song et al., 2004). Another pathway involving the genes *female*--

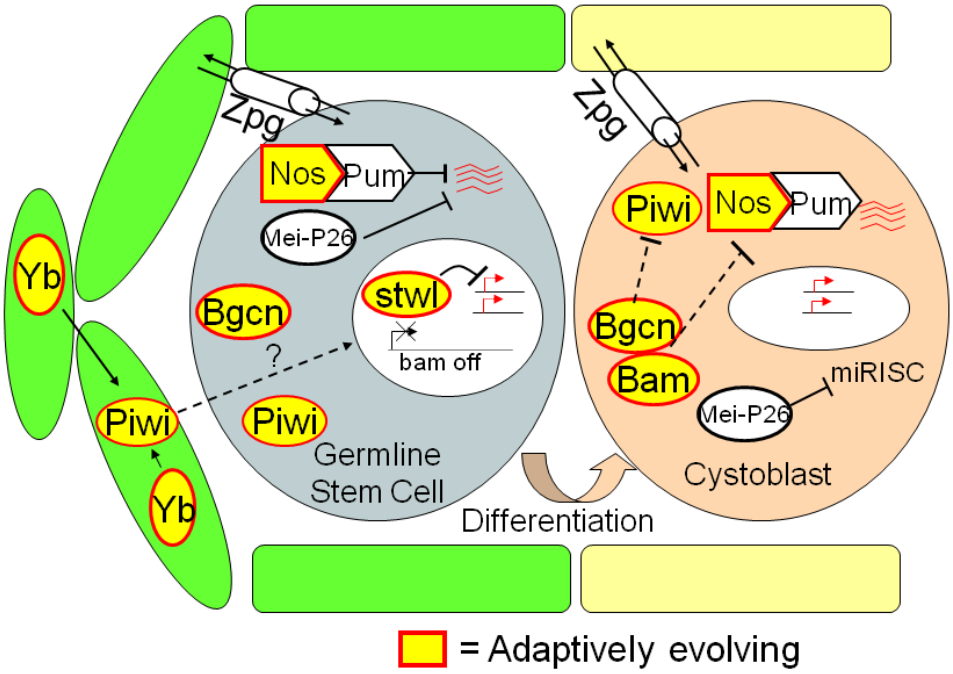


Figure 2.1: Schematic of the GSC niche with genes diverging under adaptive evolution. Figure is adapted from Wong *et. al*. (2005). We surveyed each gene for genetic variation in *D. melanogaster* and *D. simulans*, and those proteins highlighted in yellow show evidence for adaptive evolution by the McDonald-Kreitman test. The GSC (blue cell) is present in a niche environment (green cells are cap and terminal filament cells) required to maintain its stem cell state. Bam is repressed in the GSC. Only when the GSC moves away from the niche is Bam expressed and this cell starts to differentiate (tan cell). Yb is a protein involved in the maintenance of GSCs and regulating their division. Piwi acts cell non-autonomously to help in the repression of Bam in the GSC. Zpg is an adherens junction protein that functions in cell signaling. Nos and Pum act as translational repressors of genes that will promote differentiation. Mei-p26 acts in concert with the miRNA machinery to also repress transcripts, some of which are shared with Nos and Pum. Bgcn is required for Bam to cause GSCs to differentiate. Bam and Bgcn antagonize the Nos/Pum complex. The cystoblast (tan cell) will undergo four mitotic divisions. Cyc A participates in the regulation of these mitotic divisions and is not shown on this diagram.

sterile(1)Yb (*Yb*) and *P-element induced wimpy testis* (*piwi*) act to maintain stem cells by repressing *bam* expression in GSCs (King et al., 2001; Szakmary et al., 2005).

Intrinsic mechanisms within the GSC play an important role in its renewal as well, at the levels of transcription and translation. Multiple levels of transcriptional regulation occur in the GSC. Stonewall (Stwl), a chromatin-associated protein, represses genes that promote differentiation (Maines et al., 2007). Mei-P26 interacts with components of the miRNA pathway and represses transcripts that will promote differentiation (Li et al., 2012b). At the translational level, Nanos (Nos) and Pumilio (Pum) bind to mRNAs that promote differentiation and inhibit their translation (Lin and Spradling, 1997; Wang and Lin, 2004).

For a GSC daughter cell (cystoblast) that has moved away from the niche, proper regulation is also required for it to differentiate. Zpg is also required at this stage to promote cystoblast differentiation (Tazuke et al., 2002; Gilboa et al., 2003). As previously mentioned, *bam*, must be expressed in the cystoblast to promote differentiation (McKearin and Ohlstein, 1995; Ohlstein and McKearin, 1997). *bam* requires the function of another gene, *benign gonial cell neoplasm* (*bgcn*) to cause GSC differentiation (Lavoie et al., 1999; Ohlstein et al., 2000). In the cystoblast, *mei-P26* has a different role by antagonizing the miRNA pathway (Neumuller et al., 2008). Finally, the cystoblast will undergo four mitotic divisions. *bam* is thought to regulate the number of mitotic divisions, and genetic interaction assays have suggested that *bam* interacts with the cell cycle factor, *cyclin A* (*cycA*) in this process (Lilly et al., 2000).

Motivated by our evidence of strong selection acting on both *bam* and *bgcn*, we have examined DNA polymorphism in populations of *D. melanogaster* and *D. simulans* for 8 additional genes (*cycA*, *mei-P26*, *nos*, *piwi*, *pum*, *stwl*, *Yb*, *zpg*) involved in GSC regulation. We were interested in determining whether the selective pressure(s) acting on *bam* and *bgcn* was specific to these two genes or whether there were selective pressures acting on the entire system of GSC regulation. The 8 genes we tested fall into two classes: genes that genetically interact with *bam* and are likely to have shared functions and genes that have non *bam*-related roles in GSC regulation. Surprisingly, we found that 4 genes of these new 8 genes (*Yb*, *nos*, *piwi*, and *stwl*) also show a statistically significant excess of amino acid fixations in *D. melanogaster* and/or *D. simulans* in a McDonald-Kreitman test (McDonald and Kreitman, 1991). These adaptively evolving genes have various molecular functions and are expressed in a range of cell types including GSCs, cysts, and surrounding somatic cells. These unexpected findings reveal that strong evolutionary forces are acting throughout the GSC regulatory pathway.

2.2 Materials and Methods

2.2.1 Fly Stocks

When possible, African populations of *Drosophila melanogaster* and *D. simulans* were used to minimize the effects of demography in our ability to detect selection (Begun and Aquadro, 1993). In some cases, different populations were used for different genes due to availability of stocks and extracted chromosomal lines. For

the *D. melanogaster* populations, stocks were made homozygous for the X, second, or third chromosomes to eliminate heterozygosity. For *D. simulans* populations, inbred lines were used. For *stwl*, *zpg*, *piwi*, and *pum* a *D. melanogaster* population from Uganda, Africa (Pool and Aquadro, 2006) and a *D. simulans* population from Lake Kariba, Zimbabwe, Africa (Pool and Aquadro, 2006) were used. For *Yb* and *mei-P26*, a *D. melanogaster* population collected from Sengua Wildlife Research Institute in Zimbabwe, Africa (Begun and Aquadro, 1994) and a *D. simulans* population from Lake Kariba, Zimbabwe (Pool and Aquadro, 2006) were used. For *cyclin A* and *nanos*, a *D. melanogaster* population sample collected from Lake Kariba, Zimbabwe, Africa (Pool and Aquadro, 2006) and an inbred *D. simulans* population from North Carolina (Aquadro et al., 1988) were used.

2.2.2 Sequencing

Genomic DNA was extracted from approximately 20 adult flies using Purgene Core Kit A DNA isolation kits (Qiagen). PCR primers were made for each gene and used for amplification in each species. Primer sequences are listed in Table 2.1. Sanger sequencing was performed by the Cornell University Genomics Core DNA Sequencing Facility (http://cores.lifesciences.cornell.edu/brcinfo/?f=1) using ABI chemistry and 3730XL DNA Analyzers. PCR as well as internal sequencing primers were used to obtain sequence on both strands for each region. Sequences were assembled and edited using Sequencher 4.9 (Gene Codes) and aligned using MEGA 4 (Tamura et al., 2007). For *piwi*, a single 4.8 kb sequence that includes all exons was amplified. This large fragment was problematic for direct sequencing, so it was

Table 2.1 Primers used in this study

Gene	Species	Name	Sequence
cyclin A (1)	mel/sim	cycaF1	CAGTTTCCAGATCCACCAAG
cyclin A (1)	mel/sim	cycaR1	TTTAGCTTACCTCGCTCTCC
cyclin A (2)	mel/sim	cycaF2	TCTTCCAGAAGAAACATCGC
cyclin A (2)	mel/sim	cycaR2	GTATTAATATCCGGCTGCTG
nanos	mel/sim	nosF1	CAGCAACTTGGAGGGCAGTG
nanos	mel/sim	nosR1	AAACCTTCATCTGTTGCTTG
fs(1)Yb	mel	yb_sim_F1	CCTCGCTAGCCGTACATATATTAG
fs(1)Yb	mel	yb_sim_R1	GGTCAGTGGACAGTGATGAAAC
mei-P26	mel/sim	mei-p26_F2	GATGGGCTTTTGTTGAACGG
mei-P26	mel/sim	mei-p26_R2	TGCTGTTGCAGATGGTGTTG
piwi	mel/sim	piwi_F1	TTCAAAGTACTCTTTCAGTTTCC
piwi	mel/sim	piwi_R1	GTCTGGGCTAGTTTCATATATGG
pumilio (1)	mel/sim	pum_F1	CCCTACTTTCAACAGCTACAC
pumilio (1)	mel	pum_R1	CAAGCCAAGAAAGTTAACC
pumilio (1)	sim	pum_sim_R1	CAAGCCAAGAAAATTAACC
pumilio (2)	mel/sim	pum_F3	GATATTTGCTTTCCTGGAAGCC
pumilio (2)	mel/sim	pum_R3	GTCTGGGGTCTTTAGTCGG
pumilio (3)	mel	pum_F4	GGCTAAGTGGTGAATACAG
pumilio (3)	sim	pum_sim_F4	AACGTTTTAATGATAGCTTG
pumilio (3)	mel	pum_R4	GAAAATGTCACTCTGGGGAC
pumilio (3)	sim	pum_sim_R4	GAAAATGTCACTCTGGAGAC
pumilio (4)	mel	pum_F8	CATTCTCCTCTATACCTTTCC
pumilio (4)	sim	pum_sim_F8	CATTCTCTTTGATACCTCTCC
pumilio (4)	mel/sim	pum_R8	GAAGTTTCCTTTGACTGCCTG
stonewall (1)	mel/sim	stwl_F1_1	GATTGTGTGAATTGCGTTTG
stonewall (1)	mel/sim	stwl_R1_1	CTAATGGGCGATTAGTGTTAC
stonewall (2)	mel/sim	stwl_F2	CTAGCCTTATCATTTCCCTC
stonewall (2)	mel/sim	stwl_R2	CTCTTTAATCAATACTCGG
zpg	mel	zpg_F1	GTCAAACTTTACAACCGCC
zpg	sim	zpg_sim_F1	GTCAAACTTTACAAGCACC
zpg	mel	zpg_R1	GATTAAACTTGGCGTCATC
zpg	sim	zpg_sim_R1	GATTAAACTTGGTGTCATC

cloned into the pCR-BluntII-TOPO plasmid (Invitrogen). Two clones of each sample were sequenced to control for PCR errors. If there was ambiguity between the two clones, a third was sequenced and the majority nucleotide was used. The *pum* locus spans over 160 kb, so four separate products were sequenced that included most of the exons (Figure 2.2A). The *stwl* locus was amplified in two separate products that included both exons (Figure 2.2B). The *cycA* locus also amplified in two separate products that include two groups of exons in the 5' and 3'region of the gene (Figure 2.2C). For *mei-P26*, only exons 3-6 were amplified. Our results based on this region are consistent with other reports that *mei-P26* has not been subject to recurrent, positive selection (Anderson et al., 2009).

2.2.3 Polymorphism Analysis

DnaSP 5.0 (Librado and Rozas, 2009) was used to perform the following tests of neutrality: Tajima's D (Tajima, 1989) and Fay and Wu's H (Fay and Wu, 2000). *P*-values for each test were obtained using the coalescent simulator in DnaSP. Recombination estimates for each locus were obtained from sex-averaged genetic maps of *D. melanogaster* (Hey and Kliman, 2002). The following were used as estimates of recombination (recombinants/bp/generation): *zpg* r = 1.71 *10^{-8}, *stwl* r = 1.03 * 10^{-8}, *Yb* r = 2.62 * 10^{-8}, *piwi* r=2.65 * 10^{-8}, *mei-P26* r = 3.54 * 10^{-8}, *cyclin A* r = 8.76 * 10^{-9}, *pum* r = 1.01 * 10^{-9}, *nos* r=1.48*10^{-8}. These estimates of recombination were also used for *D. simulans* as an approximation of their recombination rate. An effective population size estimate of 10^6 was used for both species (Kreitman, 1983)

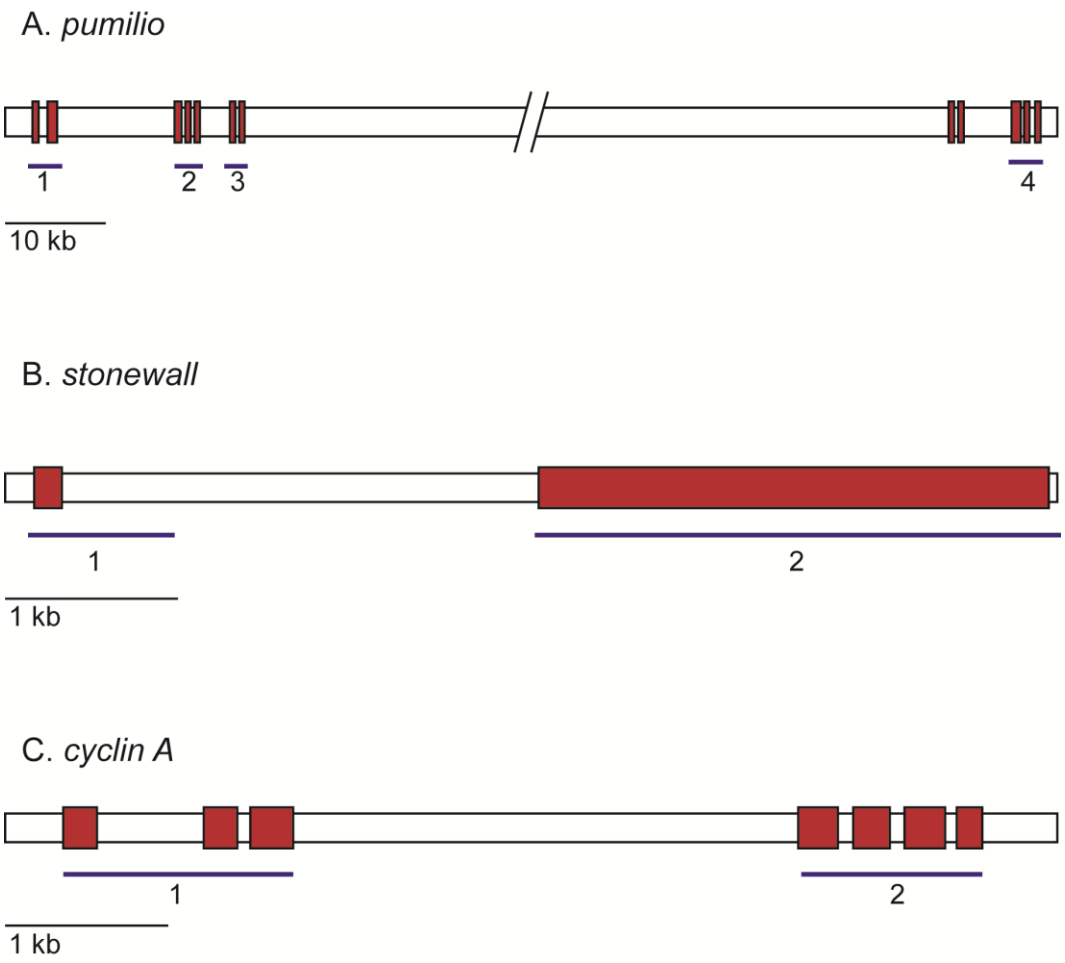


Figure 2.2: Sequenced fragments of *pumilio*, *stonewall*, and *cyclin A*.
(A) *pum* sequencing. The *pum* locus spans approximately 160 kb, so four different regions of *pum*, labeled 1-4, that include 10 of 12 exons in were individually sequenced. The diagram corresponds to the *pum*-A isoform. The center hashes denote where internal sequence (~70 kb) was removed to allow for ease of viewing. (B) *stwl* sequencing. Two fragments of *stwl* were amplified, labeled 1-2. (C) *cycA* sequencing. Two fragments of *cycA* were amplified, labeled 1-2. The blue lines denote the amplified fragments. Red boxes denote exons.

although this is likely a conservative underestimate for *D. simulans* (Aquadro et al., 1988; Moriyama and Powell, 1996).

DnaSP was also used to perform the McDonald-Kreitman test (McDonald and Kreitman, 1991), using *D. yakuba* to polarize along which lineage each fixed difference occurred. Any sites that were either missing in *D. yakuba* or different in all three species were not included in the polarized analysis.

2.2.4 Divergence Analysis

Relative rates of nonsynonymous to synonymous substitutions were estimated using Phylogenetic Analysis by Maximum Likelihood (PAML) (Yang, 1997; Yang, 2007). The PAML analyses of *nos, pum, zpg, cycA*, and *mei-P26* were included in the report by the Drosophila 12 Genomes Consortium, and the PAML results were obtained from ftp://ftp.flybase.net/12_species_analysis/ (Drosophila 12 Genomes Consortium et al., 2007). For genes not included in the Drosophila 12 Genomes Consortium analyses (*stwl, piwi*, and *Yb*) single sequences from *D. melanogaster, D. simulans, D. sechellia, D. yakuba, D. erecta*, and *D. ananassae* were used in the analysis. The models M0 vs M3, M7 vs M8, and M8 vs M8a were compared. Consistent with the Drosophila 12 Genomes Consortium analyses, each run was performed using three tree topologies: 1) *D. yakuba* and *D. erecta* as sister species 2) *D. yakuba* as an outgroup and 3) *D. erecta* as an outgroup. Each model comparison was run under three different initial ω values to assure that convergence was to a global and not local maximum.

2.3 Results

2.3.1 Polymorphism-based analyses

Gene function and sample size data from African populations of *D. melanogaster* and either African or North American *D. simulans* are reported in Table 2.2, and standard summary statistics for each gene in Table 2.3 and Table 2.4. Polymorphism data reveal significant departures from neutrality at 7 out of 16 gene/species comparisons (Table 2.3 and 2.4) in a direction that is consistent with either a recent selective sweep having occurred or recent population expansion (Tajima, 1989; Fay and Wu, 2000). The remaining genes/species have levels of polymorphism comparable to previous studies, and do not reject neutrality in tests of Tajima's D or Fay and Wu's H (Table 2.3 and 2.4).

We find that *D. simulans* levels of nucleotide variability are generally higher than those seen in *D. melanogaster,* consistent with previous results (Aquadro et al., 1988). The only exceptions were seen for the *cycA* and *mei-P26.* For *cycA* North American strains of *D. simulans* were used. Cosmopolitan strains of both *D. melanogaster* and *D. simulans* are believed to have undergone bottlenecks when the flies were migrating out of Africa, which could explain the lower variability we see (Ometto et al., 2005; Thornton and Andolfatto, 2006).

Table 2.2: Genes surveyed in this study

Gene	Function	Sample size	
		D. melanogaster	*D. simulans*
cyc A*	Regulation of cyst mitotic divisions		
Segment 1		9	10
Segment 2		9	10
Yb	GSC maintenance/division	19	1
mei-P26*	GSC maintenance	19	10
nos*	GSC maintenance	9	10
pum*	GSC maintenance		
Segment 1		17	9
Segment 2		11	10
Segment 3		19	9
Segment 4		18	7
piwi	GSC maintenance	10	6
stwl	Chromatin factor, GSC maintenance		
Segment 1		18	8
Segment 2		15	9
zpg*	GSC adherens junction	18	10

* indicates that gene has a genetic and/or physical interaction reported with *bam*. For *pumilio*, four separate regions were amplified and analyzed, labeled as 1-4. For *stonewall* and *cycA* two separate regions were amplified, labeled as 1 and 2.

Table 2.3: Nucleotide polymorphism estimates and Tajima's and Fay and Wu's tests of neutrality

Gene	Species	S	θ	π_{Tot}	π_{Syn}	π_{Non}	D	H
cycA 1	*D. melanogaster*	14	0.0051	0.0046	0.0025	0.0054	-0.4514	-1.1667
	D. simulans	10	0.0035	0.0044	0.0277	0.0000	1.2802*	2.400*
cycA 2	*D. melanogaster*	15	0.0085	0.0074	0.0157	0.0010	-0.5553	-4.333
	D. simulans	11	0.0056	0.0061	0.0164	0.0000	0.4101	1.4222
mei-P26	*D. melanogaster*	26	0.0062	0.0061	0.0181	0.0000	-0.0743	-0.6608
	D. simulans	21	0.0055	0.0031	0.0110	0.0000	-1.5532*	3.2889
nos	*D. melanogaster*	21	0.0044	0.0045	0.0090	0.0009	0.4226	-1.1389
	D. simulans	35	0.0097	0.0093	0.0150	0.0042	-0.1436	3.556
piwi	*D. melanogaster*	103	0.0079	0.0074	0.0196	0.0024	-0.1328	0.6222
	D. simulans	196	0.0222	0.0204	0.0368	0.0025	-0.7038***	11.4667
stwl 1	*D. melanogaster*	21	0.0092	0.0058	0.0000	0.0001	-1.4509**	-0.8366
	D. simulans	17	0.0088	0.0064	0.0119	0.0000	-1.3931**	-8.500**
stwl 2	*D. melanogaster*	49	0.0051	0.0048	0.0123	0.0025	-0.5371	-2.5810
	D. simulans	43	0.0053	0.0050	0.0097	0.0033	-0.3297	-5.7500
Yb	*D. melanogaster*	88	0.0079	0.0060	0.0129	0.0028	-0.9872***	-2.8947
	D. simulans	111	0.0128	0.0128	0.0259	0.0085	-0.0227	-13.861*
zpg	*D. melanogaster*	41	0.0095	0.0113	0.0413	0.0001	0.7910	-0.8100
	D. simulans	60	0.0164	0.0148	0.0286	0.0015	-0.4723	0.8890

S = segregating sites, θ = nucleotide diversity, π_{Tot} = total diversity, π_{syn} = synonymous diversity, π_{non} = nonsynonymous diversity, D = Tajima's D statistic, H = Fay and Wu's H statistic. Each amplified region of *stwl* and *cycA* was analyzed separately, see materials and methods. *P<0.025, **P<0.01, ***P<0.001

Table 2.4: Summary statistics for different *pumilio* amplicons

Gene	Species	Length (bp)	S	θ	π_{Tot}	π_{Syn}	π_{Non}	D	H
pumilio 1	*D. melanogaster*	2010	26	0.0040	0.0046	0.0033	0.0012	0.4248	1.993
	D. simulans		103	0.0202	0.0169	0.0142	0.0003	-1.2186	1.333
pumilio 2	*D. melanogaster*	874	10	0.0052	0.0040	0.0072	0.0000	-0.9925	1.545
	D. simulans		33	0.0142	0.0144	0.0346	0.0005	-0.3517	1.600
pumilio 3	*D. melanogaster*	825	10	0.0040	0.0046	0.0172	0.0020	0.4983	-0.1287
	D. simulans		72	0.0400	0.0388	0.0685	0.0014	0.0457	8.083
pumilio 4	*D. melanogaster*	2160	13	0.0021	0.0020	0.0070	0.000	-0.2597	0.02614
	D. simulans		50	0.0095	0.0089	0.0207	0.0011	-0.3854	-0.0476

S = segregating sites, D = Tajima's D statistic, H = Fay and Wu's H statistic, θ = nucleotide diversity, π_{Tot} = total diversity, π_{syn} = synonymous diversity, π_{non} = nonsynonymous diversity. See Figure 2.2 for location of each amplicon.
No tests show a significant departure from the equilibrium neutral expectations.

2.3.2 Polymorphism and Divergence-based tests

We used the McDonald-Kreitman (MK) test to determine if the ratios of nonsynonyous to synonymous polymorphisms are consistent with those for nonsynonymous to synonymous divergence, as would be expected under the standard neutral model (McDonald and Kreitman, 1991). We compared total polymorphism and divergence between *D. melanogaster* and *D. simulans* as well as lineage-specific polymorphism and divergence to determine if there is evidence of lineage-specific adaptive evolution. The MK test was polarized using the outgroup species *D. yakuba*.

The MK results revealed surprisingly that five of the eight genes surveyed had a significant MK test in at least one lineage (Table 2.5). A significant MK test can be due to deviations in any of the table's four cells. A significant excess of nonsynonymous, fixed differences due to recurrent, positive selection would be predicted provided that synonymous sites are evolving neutrally. For *Yb*, *nos*, *stwl*, and *piwi*, rejection of the MK test is consistent with an excess of nonsynonymous, fixed differences. The first three genes each have elevated values of d_N/d_S (*stwl* d_N/d_S = 0.590, *nos* d_N/d_S = 0.571, *Yb* d_N/d_S = 0.627) compared to the genome-wide average of 0.0125 (Larracuente et al., 2008), and the method of Bauer DuMont *et al*. (2004) fails to detect directional selection for synonymous codon usage in any of these genes. *piwi* shows a large number of synonymous polymorphism compared to the other genes studied. However, the method of Bauer DuMont *et al.* (2004) does not detect an significant departure from equilibrium neutrality expectations with respect to synonymous substitutions. Therefore, the rejection of neutrality for *piwi* is also likely due to an excess of nonsynonymous fixations along the *D. simulans* lineage.

Table 2.5: McDonald-Kreitman Tests for GSC genes

*P<0.05, **P<0.01, ***P<0.001

Gene	Species	Synonymous		Nonsynonymous		*P*-value
		Polymorphism	Divergence	Polymorphism	Divergence	
cyc A	Combined	17	14	4	6	0.414
	D. melanogaster	7	10	4	1	0.127
	D. simulans	10	3	1	2	0.142
mei-P26	Combined	24	44	0	0	NA
	D. melanogaster	14	22	0	0	NA
	D. simulans	10	22	0	0	NA
nos	Combined	11	11	7	23	0.075
	D. melanogaster	5	9	1	8	0.208
	D. simulans	6	2	6	15	0.033*
piwi	Combined	84	54	26	22	0.496
	D. melanogaster	32	32	13	8	0.450
	D. simulans	55	15	13	12	0.019*
pum	Combined	77	51	11	5	0.590
	D. melanogaster	19	38	8	4	0.050*
	D. simulans	58	10	5	1	1.000
stwl	Combined	45	62	48	124	0.019*
	D. melanogaster	27	27	25	47	0.101
	D. simulans	19	31	23	60	0.250
Yb	Combined	86	62	80	149	0.000012***
	D. melanogaster	42	28	29	58	0.001**
	D. simulans	46	24	52	61	0.009**
zpg	Combined	53	14	6	4	0.231
	D. melanogaster	28	10	1	2	0.200
	D. simulans	26	4	5	2	0.571

For *pum*, the MK test is marginally significant on the *D. melanogaster* lineage. However, there is also a paucity of nonsynonymous fixed differences and no evidence for selection on synonymous sites. Instead, there seems to be an excess of nonsynonymous polymorphism for *pum*. *pum* is in a region of low recombination (Hey and Kliman, 2002). Since the efficacy of selection is reduced in regions of low recombination (Haddrill et al., 2007), it is likely that an accumulation of weakly deleterious nonsynonymous polymorphisms and not positive selection for amino acid diversification is causing the MK-test departure at *pum*.

2.3.3 Divergence-based analyses

We used PAML (Yang, 1997; Yang, 2007) to test whether any of the eight GSC regulatory genes have experienced recurrent, adaptive evolution at the same subset of codons across the 12 species sequenced genomes spanning the genus Drosophila,. We find evidence of recurrent, positive selection at specific codons only for *Yb* (Table 2.6). Using both M7 versus M8 and M8 versus M8a, we find that the data fit a model of selection better than a null model. Approximately 28 of the codons are predicted to be in the selective class with a codon-specific d_N/d_S (= ω) of 2.6. However, only two codons in this class have predicted posterior probabilities greater than 0.80, and they do not fall in areas of known domains. That the other genes with significant MK test results did not reject using PAML indicates either that the positive selection is not acting on the same set of codons across species, or that the selection is acting on different codons in different lineages.

Table 2.6 PAML results

Gene	-2Δl(Model 7 vs 8)	-2Δl(Model 8 vs 8a)
cyclin A	0.306	0.345
mei-P26	0.132	0.500
nanos	0.201	0.462
piwi	0.098	0.215
pumilio	0.735	0.500
stonewall	0.806	0.368
fs(1)Yb	8.841*	6.691*
zpg	0.735	0.500

*P<0.05

2.4 Discussion

2.4.1 Multiple GSC regulatory genes show evidence of adaptive evolution

We demonstrate that seven of eight genes involved in GSC regulation reject a neutral model of evolution in at least one test and species (Tables 2.3, 2.4, 2.5). Only *zpg* failed to show evidence of any departures from an equilibrium neutral model. Five genes *(cycA*, *mei-P26*, *piwi*, *stwl*, and *Yb*) reject a site-frequency-spectrum based test in a manner consistent with either a recent selective sweep or population expansion. *piwi*, *stwl* and *Yb* also reject a MK test, providing suggesting that the recent sweeps are just the lastest of many recurrent selective fixations of nonsynonymous substitutions. Additionally, *nos* also rejects a MK test due to an excess of nonsynonymous, fixed differences similar to what was seen for *bam* and *bgcn*. Three of the eight new genes examined have no known interaction or dependence on *bam* function (*stwl*, *piwi,* and *Yb*), while one gene (*nos*) interacts with *bam* (Chen and McKearin, 2005; Li et al., 2009).

Our hope was that the identification of additional genes divergine under adaptive evolution would provide insight into the selective pressures acting on *bam* and *bgcn*. Instead, the pervasive evidence for natural selection we observe and the diverse functions and expression patterns of these genes suggest that there are likely multiple selective pressures acting on genes important in GSC regulation. For example, *Yb* is expressed in the stem cell niche (King and Lin, 1999; King et al., 2001), while *stwl* is present in GSCs where it binds chromatin (Clark and McKearin, 1996; Maines et al., 2007), making it less likely that the same specific selective pressures act on both. Additionally, recent studies have identified novel roles for *piwi*

and *Yb* in the germline and somatic repression of transposable elements (Aravin et al., 2007; Olivieri et al., 2010; Saito et al., 2010). Below, we discuss possible selective pressures that could be acting on genes involved in GSC regulation.

2.4.2 Selective pressures acting on genes involved in GSC regulation

Several mechanistic and evolutionary hypotheses have been proposed to explain the evolutionary causes of positive selection inferred for *bam* and *bgcn*. Some of these selective pressures may also drive the adaptive evolution of other genes involved in GSC regulation.

Civetta *et. al*. (2006) proposed that species-specific changes in rates of proteolysis could drive protein sequence divergence. This proposal was supported by the observation that *bam*'s expression is transient and by previous studies in *C. elegans* that have shown that transiently expressed genes have elevated rates of protein evolution (Cutter and Ward, 2005). While this could influence the molecular evolution of *bam,* and potentially *bgcn* which is also transiently expressed (Ohlstein et al., 2000), it is unlikely to explain all selection acting on GSC gene evolution since *piwi, Yb, stwl,* and *nanos* have much broader patterns and timings of expression (Clark and McKearin, 1996; Forbes and Lehmann, 1998; Cox et al., 2000; Szakmary et al., 2009).

We had previously hypothesized that coevolution with external pathogens infecting the germline could underlie the elevated nonsynonymous *bam* and *bgcn* divergence along the *D. melanogaster* and *D. simulans* lineages (Bauer DuMont et al., 2007). Two maternally-inherited bacterial endosymbionts (*Wolbachia* and

Spiroplasma) have been detected in some but not all species of Drosophila (Mateos et al., 2006). Infection by *Wolbachia* can have beneficial effects in some species by increasing resistance to viral infections, which may explain their widespread presence (Hedges et al., 2008; Teixeira et al., 2008). However, *Wolbachia* infection can also reduce fecundity due to cytoplasmic incompatibilities in crosses between infected and uninfected individuals (Fry et al., 2004). There is likely to be a delicate balance in controlling endosymbiont proliferation within a cell so that the host can receive benefits from the endosymbiont but minimize any deleterious effects. We hypothesized that a selective pressure towards decreasing the detrimental effects of *Wolbachia* or *Spiroplasma*, for example on postfertilization incompatibility, could drive an "arms race" between GSC regulatory genes and endosymbionts (Bauer DuMont et al., 2007). This arms race within each infected species would drive independent protein sequence evolution at these genes in each lineage, and thus could potentially explain the observed high levels of protein divergence.

While we still believe endosymbionts provide a plausible selective pressure for some GSC regulatory genes (such as *bam*, *bgcn*), the expression patterns and known functions of other GSC genes (such as *Yb*, *piwi*, *stwl*) suggest that different pressures may be acting on GSC regulatory genes as well (Clark and McKearin, 1996; Aravin et al., 2007; Brennecke et al., 2008). One possibility is intracellular parasites such as transposons and viruses. Transposons are selfish genetic elements that can propagate throughout the genome, resulting in deleterious effects on their host. Recent studies demonstrated that many taxa, including Drosophila, have a small RNA silencing pathway, termed the piRNA pathway, that is active in the germline and provides an

adaptive defense against transposons (reviewed in Aravin et al., 2007). Many piRNA pathway genes have also been shown to adaptively evolve (Obbard et al., 2009; Kolaczkowski et al., 2011). *piwi* and *Yb* are required for the proper silencing of transposons (Aravin et al., 2007; Olivieri et al., 2010; Saito et al., 2010). Therefore, the adaptive evolution seen in these two proteins may reflect their involvement in silencing transposons. This hypothesis has been suggested for *piwi* as other studies have shown it is adaptively evolving (Obbard et al., 2009; Kolaczkowski et al., 2011). We suggest it can potentially explain the positive selection observed at *Yb* as well. Additionally, it is possible that selective pressure to repress transposons may be driving the adaptive evolution of *stwl* as well since some chromatin-associated proteins are involved in transposon silencing (Klattenhoff et al., 2009; Rangan et al., 2011).

Species-specific changes in life history and the timing of reproduction could also pose changing selective pressures on the germline (Schmidt and Paaby, 2008), though our limited knowledge of the ages of reproduction for natural populations of Drosophila limits our ability to test this hypothesis. Some alternative hypotheses such as sexual selection and sexual conflict (Swanson and Vacquier, 2002) cannot be formally rejected but seem implausible. For example, most theories of sexual selection predict strong effects on pre-mating traits, which are highly unlikely to be influenced by GSC regulatory genes such as *bam* and *bgcn* that have restricted expression patterns in pre-gametic germline tissues. Likewise, sexual conflict, whereby one sex manipulates the reproductive fitness of the other sex, is much more

likely to occur for molecules that are transmitted between males and females, a function that is implausible for any of the GSC regulatory genes in this study.

In the future it will be important to test whether these positively selected GSC genes function in the specific biological processes that we hypothesize are driving their adaptive evolution. For example, does *stonewall* play a role in the repression of transposons or do *bam* or *bgcn* play a role in regulating the transmission of bacterial endosymbionts? Although showing these genes play a role in these processes does not prove that these selective pressures have driven the adaptive evolution of these genes, it would provide support to these hypotheses. We present evidence of both recurrent, adaptive evolution as well as evidence consistent with recent selective sweeps occurring in *piwi*, *stwl*, and *Yb*. While we may not know the selective pressure acting on these genes, this data suggests the selective pressure may still be acting on these genes. Additional insight may come from sampling these genes from additional Drosophila species to determine whether this has been a long-term selective pressure across many Drosophila or whether it is specific to *D. melanogaster* and *D. simulans*.

CHAPTER 3

SELECTIVE PRESSURES IN THE FEMALE GERMLINE HAVE DRIVEN THE ADAPTIVE EVOLUTION OF THE *BAG OF MARBLES* GENE

3.1 Introduction

Population genetic and comparative analyses in diverse taxa have shown that many genes involved in reproduction are evolving under adaptive evolution (Swanson and Vacquier, 2002; reviewed in Panhuis et al., 2006; Turner and Hoekstra, 2008). Many selective pressures have been hypothesized to drive the adaptive evolution of those reproductive genes including sexual conflict, sexual selection, pathogen resistance, and avoidance of interspecific fertilization (Swanson and Vacquier, 2002; Clark et al., 2006; Haerty et al., 2007). While population genetic and comparative approaches have been valuable in identifying adaptively evolving genes (Swanson et al., 2004; Begun et al., 2007; Drosophila 12 Genomes Consortium et al., 2007; Haerty et al., 2007; Kosiol et al., 2008; Larracuente et al., 2008; Dean et al., 2009), a combination of these approaches with functional analysis is needed to identify the adaptive phenotypes and to determine the contribution of these selective pressures.

A more recent focus of evolutionary biology has been to determine the functional consequences of adaptive evolution (MacCallum and Hill, 2006; Jensen et al., 2007). While this can provide exciting insight into the phenotypes under selection it can often be a difficult task, as adaptively evolving proteins may not be well defined in terms of their function or the pathways in which they participate. We have

previously shown that the *bag of marbles* (*bam*) gene has experienced recurrent, adaptive evolution in *D. melanogaster* and *D. simulans* (Civetta et al., 2006; Bauer DuMont et al., 2007). We observe more than 4 fold the expected number of amino acid substitutions between these two species, and the substitutions are spread throughout the entire Bam protein length. *bam* is involved in the early stages of reproduction, as it regulates germline stem cell differentiation and germline cyst development in both males and females.

Germline stem cells (GSCs) are present in a niche environment that is required to maintain their stem cell state (reviewed in Wong et al., 2005; Xie et al., 2008). When a stem cell asymmetrically divides, the daughter cell, a cystoblast, moves away from the niche which relieves repressive mechanisms, allowing for the daughter cell to differentiate. The cystoblast will undergo four mitotic divisions in synchrony to generate an interconnected, 16-cell cyst. In females, one of these cells will become the oocyte and enter meiosis while the remaining 15 nurse cells will polyploidize and provide nutrients to the oocyte. In males, all 16 cells will enter meiosis and give rise to mature sperm.

In females, *bam* is the key factor for inducing GSCs to differentiation. Therefore, *bam* is transcriptionally repressed in the GSC (Chen and McKearin, 2003a; Chen and McKearin, 2003b; Song et al., 2004). In the cystoblast, repression of *bam* is relieved, leading to its expression and the start of the cascade to promote differentiation. Bam expression is transient, as its protein is present only in late cystoblasts, and 2, 4, and 8 cell cysts (McKearin and Ohlstein, 1995). The molecular function of *bam* is not fully understood, but Bam physically interacts with and requires

the function of *benign gonial cell neoplasm* (*bgcn*) in GSC differentiation (Lavoie et al., 1999; Ohlstein et al., 2000; Li et al., 2009; Shen et al., 2009). Bgcn is a member of the DExH-box family of ATP-dependent RNA helicases. Bgcn lacks the sites needed for ATP binding and helicase activity, but it is believed to be able to bind RNA which has led to the hypothesis that Bam and Bgcn act together as translational repressors (Ohlstein et al., 2000).

In males, *bam* is not required for the induction GSC differentiation, as *bam* mutant GSCs differentiate but continue undergoing mitotic divisions and never enter meiosis (McKearin and Spradling, 1990; Gonczy et al., 1997; Insco et al., 2009). As in females, Bam protein is expressed transiently in males as it is present in only 4, 8, and 16 cell cysts (Insco et al., 2009).

Bam also plays roles downstream of GSC differentiation in both males and females. Bam localizes to the fusome, an ER-like organelle that interconnects the cells of a cyst, mediates the synchrony of the mitotic divisions, and likely determines the future oocyte (McKearin and Ohlstein, 1995; de Cuevas and Spradling, 1998). This localization requires Bgcn (Lavoie et al., 1999), and *bam* mutants show a reduction in fusome vesicles (McKearin and Ohlstein, 1995). Bam's expression in mitotically-active cysts has led to the hypothesis that it plays a role in counting cyst divisions. While there are some results consistent with this role in females (McKearin and Ohlstein, 1995; Hawkins et al., 1996; Lilly et al., 2000), in males this role has been described in more detail. Insco et. al. (2009) have shown that the accumulation of Bam to a critical threshold is required for cysts to cease mitotic divisions and initiate spermatocyte differentiation.

To identify the functional consequences of *bam*'s divergence, we have developed a transgenic system to assay the ability of a *bam* ortholog from *D. melanogaster* or *D. simulans* to rescue the female and male sterility of a *D. melanogaster bam* mutant. Using this assay, we find that while *D. simulans bam* can rescue *D. melanogaster bam* mutant male sterility, it fails to fully rescue *D. melanogaster bam* mutant female sterility. Our results suggest that the selective pressure driving the adaptive evolution of *bam* is in the female germline, and we discuss possible selective pressures.

3.2 Materials and Methods

3.2.1 Drosophila stocks

All stocks were cultured at room temperature on standard yeast-glucose media. The $bam^{\Delta 86}$, bam^{BW}, bam^{BG}, and $bgcn^{1}$ stocks are described in FlyBase. The $bam^{\Delta 59}$ allele was generated through a P-element excision of bam^{1} (D. McKearin, pers. comm.). We sequenced this allele and discovered that the excision deleted all but the 31 amino acids from the C-terminal end of the protein. All five stocks were kindly provided by Dr. Dennis McKearin (HHMI).

3.2.2 DNA constructs

D. melanogaster bam transgene construction

To generate a *D. melanogaster bam* (abbreviated as *mel-bam*) transgene driven by its native promoter, we amplified a 4.1 kb fragment from genomic DNA of the sequenced *D. melanogaster* strain, *y*; *cn bw*; *sp*, using primers 904 and 905. This

fragment contains approximately 1.7 kb upstream of the *bam* start codon and approximately 1 kb downstream of the stop codon. The PCR product was cloned into the pCR-Blunt II-TOPO (Invitrogen) vector to generate the plasmid p{mel-bam}. The insert was sequenced on both strands and was identical to the reference sequence.

A three-piece fusion PCR strategy was used to incorporate a Yellow fluorescent protein (YFP) tag into the *bam* coding region at the C-terminus (abbreviated as *mel-bam-yfp*). Two products were amplified using p{mel-bam} as the template with the primer pairs 906/907 and 908/909. These products correspond to parts of the *D. melanogaster bam* sequence directly upstream and downstream of the native stop codon. The third product containing the YFP tag was amplified using p{w^{+mC} UAS-Lhr::Venus=UAS-Lhr::YFP} as the template (Brideau et al., 2006) with primer pair 910/911. All three products were gel-purified and used as templates for fusion PCR. The templates were allowed to prime one another for 6 cycles, and then primer pair 906/909 was added to amplify the final product. The final product was cloned into pCR-Blunt II-TOPO and sequenced. The insert was subcloned into p{mel-bam} using *Nde*I and *Stu*I restriction enzymes, generating p{mel-bam-yfp}. The full-length and tagged *D. melanogaster bam* transgene was subcloned in the transformation vector pCasper4\attB (Maheshwari and Barbash, 2012) using *Not*I and *Kpn*I restriction enzymes and the insert was sequenced entirely.

Table 3.1 Primers used in this study

No.	Sequence
661	TTAGCTTCTGAAGCGAGGTACAC
662	CACCATGCTTAATGCACGTGACATG
844	CCGCTTCAAGGGACAGTATC
845	GACAATCTCCTTGCGCTTCT
904	GCAAAAGATCTTCTGCACCCTCTG
905	TTTAGCGGATTCACAAGGGATCTC
891	TTTAGCGGATTCACAAGTGATCTC
906	CTGCATATGATTGGTCTGCACGG
907	AGCTCCTCGCCCTTGCTCACCATGCTTCTGAAGCGAGGTACACGTCC
908	TCTCGGCATGGACGAGCTGTACAAGTAAACTAATGCTGTGCACATCGATA
909	CCAGAAAGATCTCAGCGAGAACATG
910	GGACGTGTACCTCGCTTCAGAAGCATGGTGAGCAAGGGCGAGGAGCT
911	TATCGATGTGCACAGCATTAGTTTACTTGTACAGCTCGTCCATGCCGAGA
926	CTCACTGTCCAATGTTCCTTC
927	CCCATGTCACGTGCATTAAGCATTATTCTTAAGTTAAATCACACAAATC
928	GATTTGTGTGATTTAACTTAAGAATAATGCTTAATGCACGTGACATGGG
929	GCTCCTCGCCCTTGCTCACCATGCTTCTGAAGCGAGGTACACGTATGG
930	CCATACGTGTACCTCGCTTCAGAAGCATGGTGAGCAAGGGCGAGGAGC
931	GCTGGAAAATCTGTTCAACGG
949	GTCGACGATGTAGGTCACGGTC
1125	CGCCTTGTCCAGTCCAAAAAG
1169	GCCCATAACTATTGAGAAACTGC
1170	GATCATGCAGGGATCTGAACAG
1480	CTGCTCCATGCTCACTGCGCCAAGCTTCTGTGACCCGCAAATGGCGAC
1481	GAGGAAGTGCCATCATCGCCACCTCGTCACACCATACGTGTACCTCGC
1479	GTCGCCATTTGCGGGTCACAGAAGCTTGGCGCAGTGAGCATGGAGCAG
1482	GCGAGGTACACGTATGGTGTGACGAGGTGGCGATGATGGCACTTCCTC

D. simulans bam transgene construction

To generate a *D. simulans bam* transgene driven by the *D. melanogaster* native promoter, we amplified the region orthologous to the *D. melanogaster* region described above from *D. simulans* w^{501} genomic DNA using the primer pair 904/891. The PCR product was cloned into the pCR-Blunt II-TOPO vector and sequenced completely. A three-piece fusion PCR strategy was used to incorporate both the *D. melanogaster* regulatory region and YFP tag simultaneously. Two products for fusion were amplified using p{mel-bam-yfp} as template with primer pairs 926/927 and 930/931, corresponding to the *D. melanogaster* 5' region and 3' regulatory region including the YFP tag, respectively. The third product for fusion was amplified from p{sim-bam} using primer pair 928/929. The gel-purified products were used as templates for fusion PCR as described above using primers 926 and 931 to amplify the final product. The fusion product was cloned into pCR-Blunt II-TOPO and sequenced. The insert was subcloned into p{mel-bam-yfp} using *Mfe*I and *Stu*I, generating p{melrsimc-bam-yfp}(abbreviated below as *sim-bam-yfp*). The full-length insert was then cloned into the *Not*I and *Kpn*I sites of pCasper4\attB, and the insert was completely sequenced.

D. simulans bam containing *D. melanogaster* PEST domain transgene construction

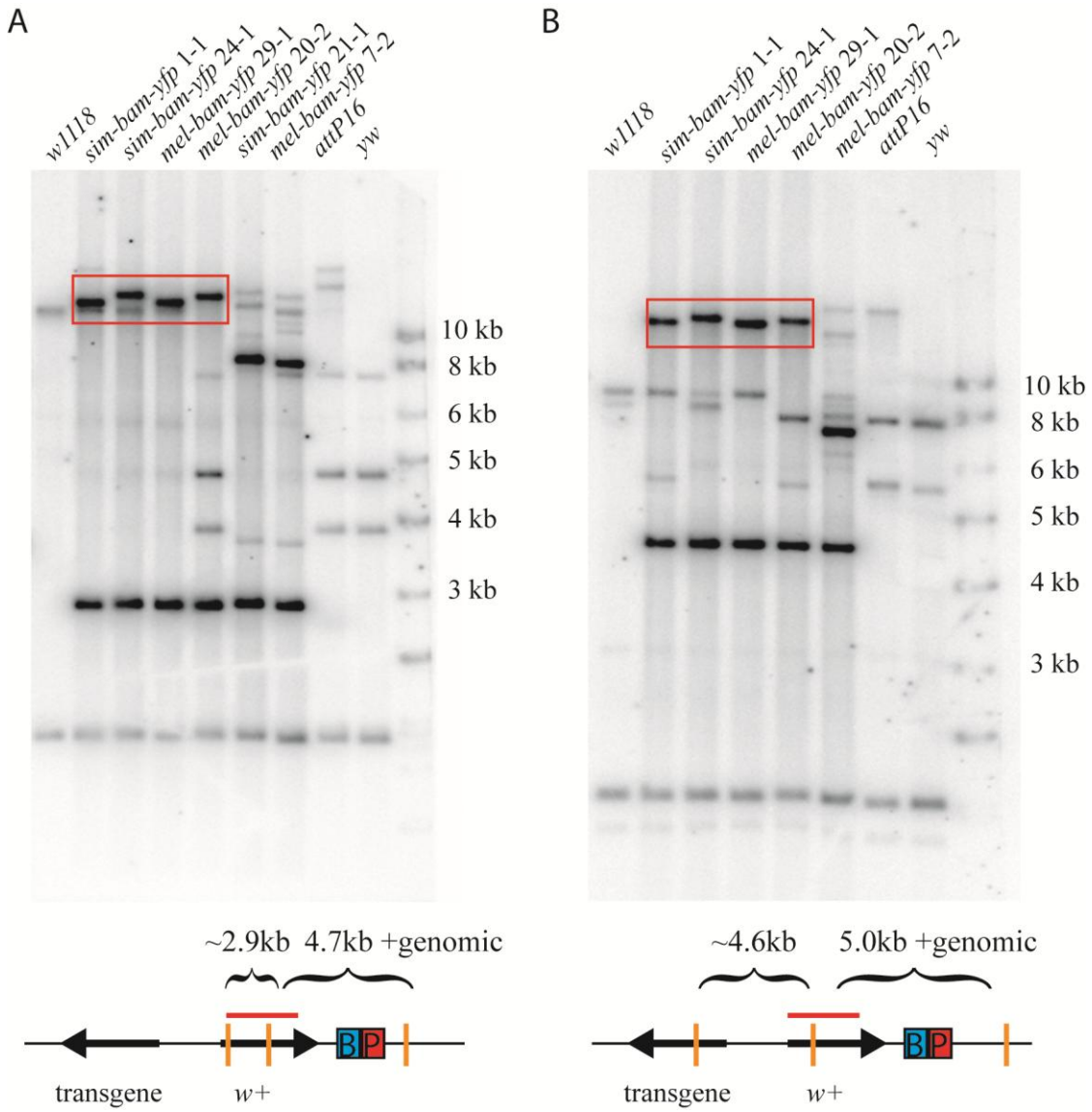
To construct a *D. simulans bam* transgene containing a *D. melanogaster bam* PEST domain, a three-piece fusion PCR strategy was used. Two fusion products were amplified using p{melrsimc-bam-yfp} with primers 926/1480 and 1481/931. The third fusion product was amplified from p{mel-bam-yfp} using primers 1479/1482.

The gel-purified products were used in fusion PCR as described above with primers 926/931 to amplify the final product. The PCR product was cloned into pCR-Blunt II-TOPO, sequenced completely, and subcloned into p{melrsimc-bam-yfp} using *Mfe*I and *Stu*I, generating p{melrsimc-melpest-bam-yfp}. The full-length construct was cloned into the *Not*I and *Kpn*I sites in pCasper4\attB, and the insert was fully sequenced.

3.2.3 Transgenic fly lines

ΦC31-mediated transformation was used to generate transformants in *D. melanogaster* (Groth et al., 2004) and was performed by Genetic Services, Inc. Correct integration was assayed using a PCR-based assay developed by Venken et al. (2006), while docking site-specificity was assayed using primers designed to specific integration sites. For the attP40 site at cytological position 25C6, the primer pair 949/1125 was used to check docking-site specificity. We discovered that the attP16 stock contains at least two attP docking sites at unknown locations. Southern blots were used to determine that p{mel-bam-yfp} and p{sim-bam-yfp} both integrated into the same *attP* site (Figure 3.1). The Southern blot probe was designed to the *white* locus present on pCasper4\attB and was generated by digesting pCasper4\attB with *DraI* and gel-purifying product the ~4.2 kb product including *w*+. We will refer to the two attP sites in attP16 as attP16a and attP16b. All transformants were then outcrossed for at least six generations to a *yw* strain that had been inbred for 10 generations, to make the genetic backgrounds similar.

Figure 3.1: Southern blot identifies transgenes in the same insertion site.
(A) Genomic DNA was digested with *EcoRV*. (B) Genomic DNA was digested with *ClaI*. Blots were incubated with a probe designed to *w+* on pCasper4\attB. Below each blot is a schematic showing the location of the *w+* probe (red bar), the restriction enzyme sites (orange), the location of the *attB* and *attP* sequences (boxes with B and P), and the approximate sizes of the digested fragments. The red box over the membrane highlights the diagnostic fragment used to determine shared integration sites. Lines *mel-bam-yfp* 29-1 and 20-2 as well as *sim-bam-yfp* lines 24-1 and 1-1 were all derived from integrations in the attP16 stock carrying multiple *attP* sites. Lines *mel-bam-yfp* 7-2 *and sim-bam-yfp* 21-1 were derived from integrations into attP40 in which only one *attP* site is present. Also run on the gels were the undocked attP16 line and *yw* and w^{1118} in which the transgenic stocks had been crossed. These data show that *sim-bam-yfp* line 1-1 and *mel-bam-yfp* line 29-1 are integrated in the same *attP* site, termed attP16a, and that *sim-bam-yfp* 24-1 and *mel-bam-yfp* 20-2 are both in a distinct site termed attP16b. These data also confirm that *mel-bam-yfp* 7-2 and *sim-bam-yfp* 21-1 are in the same insertion site.


A
w1118
sim-bam-yfp 1-1
sim-bam-yfp 24-1
mel-bam-yfp 29-1
mel-bam-yfp 20-2
sim-bam-yfp 21-1
mel-bam-yfp 7-2
attP16
yw
10 kb
8 kb
6 kb
5 kb
4 kb
3 kb
~2.9kb
4.7kb +genomic
transgene
w+
B
w1118
sim-bam-yfp 1-1
sim-bam-yfp 24-1
mel-bam-yfp 29-1
mel-bam-yfp 20-2
mel-bam-yfp 7-2
attP16
yw
10 kb
8 kb
6 kb
5 kb
4 kb
3 kb
~4.6kb
5.0kb +genomic
transgene
w+

3.2.4 Fertility assays

All crosses were performed at room temperature (22-23°C). Prior to crossing all flies were aged for 2-3 days post-eclosion on media supplemented with yeast. In female fertility experiments, single transgenic females were crossed to two wildtype Canton S (CS) males. The trio of flies were transferred to a new vial every five days for a total of 15 days and then discarded. Progeny from each vial were counted for 8 days after the first flies eclosed. In male fertility experiments, single males were mated to two wildtype CS females as described above. In sperm exhaustion assays, single males were mated to two wildtype CS females. The males were aspirated without anesthetizing into new vials containing two fresh CS females every day for 5 days. The females remaining in the vial were transferred to a new vial every five days for 10 days, and fertility was assessed by scoring the number of progeny that eclosed over 8 days.

For female fertility assays, the transgenes were crossed into the *bam* mutant background $bam^{\Delta 86}/bam^{\Delta 59}$. We found that *mel-bam-yfp;bam*$^{-}$ females were as fertile as *D. melanogaster bam* heterozygous siblings ($bam^{\Delta 59}$/+) and this result did not change when the comparison of *mel-bam-yfp;bam*$^{-}$ versus the *D. melanogaster bam* heterozygote (bam^{BG} /+) was made in a different genetic background, $bam^{\Delta 86}/bam^{BG}$.

For male fertility we found that use of $bam^{\Delta 86}/bam^{\Delta 59}$ resulted in reduced fertility of *mel-bam-yfp* relative to the *D. melanogaster bam* heterozygous control, suggesting that background effects in these mutants reduce male fertility. It is also likely that combinations of $bam^{\Delta 86}$ or $bam^{\Delta 59}$ with bam^{1}, the chromosome from which they were derived, will share these background effects. Therefore, all male fertility

experiments were done with the transheterozygous combination, $bam^{\Delta 86}/bam^{BG}$ which are independently-derived mutations of *bam* and where there is no reduction of fertility of *mel-bam-yfp*;bam^{-}, relative to the *D. melanogaster bam* heterozygote.

3.2.5 Quantitative RT-PCR

Flies were aged 2 days on media supplemented with yeast. Ovaries or testes were dissected in 1XPBS, and total RNA was isolated from 10 ovaries or ~100 testes using Trizol reagent (Invitrogen) following the manufacturer's protocol. Samples were treated with 20 units DnaseI at 37°C for 2 hours (Roche) and purified using RNeasy columns (Qiagen) following the manufacture's protocol. cDNA was generated from 2ug of ovarian or 5ug of testes total RNA using the Superscript III First Strand Synthesis kit (Invitrogen) and oligo-dT primers following the manufacturer's protocol. Quantitative RT-PCR was performed on a Biorad MyiQ cycler using iQ SYBR Green Supermix (Biorad). For *bam*, primer pair 1160/1170 amplified *bam* from both species with high efficiencies. For *rpl32*, primer pair 844/845 from Maheshwari and Barbash (2012) was used. The standard curve method was used to estimate *bam* and *rpl32* levels. Three technical replicates were performed from at least three biological replicates for each sample.

3.2.6 Co-Immunoprecipitation Experiments

D. simulans bam was amplified from w^{501} ovarian cDNA using primers 662/661, cloned into pENTR/D-TOPO vector (Invitrogen), verified by sequencing, and recombined into destination vectors using LR-Clonase II (Invitrogen) following

manufacturer's directions. *D. simulans bam* was recombined into pAFHW containing both Flag and HA epitope tags (http://emb.carnegiescience.edu/labs/murphy/Gateway%20vectors.html). *D. melanogaster bam* in pAFHW and *D. melanogaster bgcn* in pAFMW were kindly provided by D. McKearin (Li et al, 2009).

Combinations of pAFMW-Bam and pAFHW-Bgcn or empty vectors were co-transfected into Drosophila S2 cells, cells incubated for 3 days, and then lysed in lysis buffer (50mM Tris-HCl pH7.8, 150mM NaCl, 0.1%NP-40). Anti-HA (Roche, 3F10) or anti-Myc (Roche, 9E10) antibodies were conjugated to 50 ul of Protein G Dynabeads (Invitrogen) in 200ul of PBST (0.01% Tween 20) at 4°C overnight with rotation. Antibody-conjugated beads were then added to cell lysate (80ug total protein) in 200ul in lysis buffer containing 1X protease inhibitor (Roche) and 1mM PMSF and incubated at 4°C overnight. Washes were performed following manufacturer's directions and Dynabeads were boiled in 1X SDS sample buffer to elute protein.

3.2.7 Western blotting

25-35 ovaries from females aged 2-3 days on media supplemented with yeast were homogenized in lysis buffer (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 1.25% TritonX-100, 1X protease inhibitor, Roche) and centrifuged at 14000 rpm at 4°C for 5 minutes. Total protein in the supernatant was estimated using the Bradford assay (Biorad) and samples were boiled in an equal volume of 4X SDS sample buffer for 5 minutes. 10-20 ug were loaded on 10% SDS-PAGE gels. Primary antibodies were anti-GFP Jl-8 (Clontech, 1:2000) and mouse anti-Tubulin T5168 (Sigma; 1:120,000).

Secondary antibodies were HRP conjugated goat anti-mouse secondary antibodies (Jackson; 1:1,000 for anti-GFP and 1:60,000 for anti-Tubulin) and were detected with ECL Western blotting substrate (Pierce).

Western quantifications were performed as described in (Shibata et al., 2011). Westerns were quantified using the Gel Analysis software in Image J (http://rsbweb.nih.gov/ij/). Standard estimates for Tubulin were plotted (x-axis) against standard estimates for YFP (y-axis) and a linear regression was used to estimate the amount of Bam-YFP present in each lane for the amount of Tubulin present in each experimental sample.

3.2.8 Immunostaining

Immunostaining was performed as in Aruna et al. (2009). For primary antibodies, the following antibodies were used: anti-GFP (Invitrogen A6544, 1:200), anti-Vasa (DSHB, 1:25), anti-1B1 (DSHB, 1:4), Polyclonal anti-Bam (1:500), monoclonal anti-Bam (1:100). Anti-Bam antibodies were provided by D. McKearin. Secondary antibodies including goat anti-rat, anti-rabbit, or anti-mouse were conjugated with Alexa fluor dyes (Molecular Probes, 1:200-1:500). Samples were mounted in Vectashield containing DAPI (Vector Laboratories) and analyzed using the Leica SP2 confocal microscope at the Cornell University Core Life Sciences Microscopy and Imaging Facility using. Images were resized in Photoshop (Adobe, version 11.0).

3.3 Results

3.3.1 A transgenic system for assessing functional effects of Bam divergence

We generated strains of *D. melanogaster* containing transgenic copies of either *D. melanogaster bam* (*mel-bam-yfp*) or *D. simulans bam* (*sim-bam-yfp*). Each *bam* ortholog was C-terminally tagged with Yellow fluorescent protein (YFP) and driven by the native *D. melanogaster* regulatory region which has been previously defined (McKearin and Spradling, 1990; Chen and McKearin, 2003b). This approach was designed in an effort to attribute any phenotypic differences to coding sequence divergence (Figure 3.2). Two integrations of each transgene were made separately in the same positions of the *D. melanogaster* genome (attP16a or attP40 on chromosome 2), and then crossed into a *D. melanogaster bam* transheterozygous, null mutant background. Initial experiments demonstrated that *mel-bam-yfp* is functional because it rescued the sterility of both *bam* mutant females and males. The nomenclature used throughout this study is described in Table 3.2.

qRT-PCR analyses from ovarian cDNA provided unexpected results (Figure 3.3). First, *bam* expression levels in *mel-bam-yfp;bam*$^{-}$ ovaries are over 10-fold less than a single *D. melanogaster bam* allele (*bam* heterozygote of *bam*$^{\Delta 59}$/+) generated from the same cross. Second, although the *sim-bam-yfp* transgene is driven by the same *D. melanogaster bam* regulatory region as the *mel-bam-yfp* transgene, the *bam* expression level in *sim-bam-yfp;bam*$^{-}$ is not significantly different from that of the *D. melanogaster bam* heterozygote. To determine if the unexpectedly low *bam* expression in *mel-bam-yfp;bam*$^{-}$ is due to a mutation caused during transformation or to a background effect, additional qRT-PCR was performed in which we found that

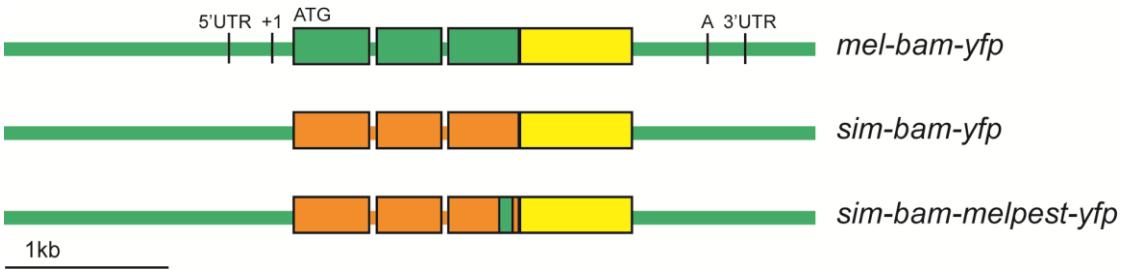


Figure 3.2: *bam* transgenic constructs.
All constructs are drawn to scale and contain the entire *bam* open reading frame (thick bars), 2 small introns and regulatory region (thin bars). Green color corresponds to *D. melanogaster* sequences, orange corresponds to *D. simulans* sequences and yellow corresponds to YFP coding sequence. ATG denotes the start codon. 5' and 3' UTR sequence boundaries are shown (Flybase). The transcription start sites is denoted as +1 (Chen and McKearin, 2003b). The poly(A) addition sequence is denoted as A (McKearin and Spradling, 1990).

Table 3.2 Nomenclature

Nomenclature	Genotype	Description
D. melanogaster bam heterozygote	Females: $bam^{\Delta 59}$/+ Males: bam^{BG}/+	*D. melanogaster* with only a single wildtype copy of *bam*
mel-bam-yfp; bam^{-} or *sim-bam-yfp*; bam^{-}	Females: *w*; φ{w^{+}; *transgene*}/+; $bam^{\Delta 86}/bam^{\Delta 59}$ Males: *w*; φ{w^{+}; *transgene*}/+; $bam^{\Delta 86}/bam^{BG}$	A single copy of a transgene in a *D. melanogaster bam* null mutant background
mel-bam-yfp; bam^{-}/+ or *sim-bam-yfp*; bam^{-}/+	Females: *w*; φ{w^{+}; *transgene*}/+; $bam^{\Delta 59}$/+	A single copy of a transgene in a *D. melanogaster bam* heterozygous background
2x *mel-bam-yfp*; bam^{-} or 2x *sim-bam-yfp*; bam^{-}	Females: *w*; φ{w^{+}; *transgene*}, φ{w^{+}; *transgene*}, /+ +; $bam^{\Delta 86}/bam^{\Delta 59}$ Males: *w*; φ{w^{+}; *transgene*}, φ{w^{+}; *transgene*}/+ +; $bam^{\Delta 86}/bam^{BG}$	Two copies of a transgene in a *D. melanogaster bam* null background. The chromosome carrying 2 copies of a transgene was generated by recombining attP40 and attP16a transgene-containing sites onto a single chromosome.
2x *sim-bam-yfp*; bam^{-}/+	Females: *w*; φ{w^{+}; *transgene*}, φ{w^{+}; *transgene*}/+; $bam^{\Delta 59}$/+	Two copies of *sim-bam-yfp* in a *D. melanogaster bam* heterozygous background. These flies are siblings of 2x *sim-bam-yfp*; bam^{-}flies.
4x *mel-bam-yfp*; +/+ or 4x *sim-bam-yfp*; +/+	Females: *w*; φ{w^{+}; *transgene*}, φ{w^{+}; *transgene*} / φ{w^{+}; *transgene*}, φ{w^{+}; *transgene*}; +/+	Four copies of *sim-bam-yfp* in a *D. melanogaster* fly wildtype for *bam*.
bam-α;bam^{-}	Females: *w; P{ry+, bam-α},*$bam^{\Delta 86}/bam^{\Delta 59}$ Males: *w; P{ry+, bam-α},*$bam^{\Delta 86}/bam^{BG}$	A *bam* transgene from Dennis McKearin's lab, stock DM25. A single copy of the transgene in a *bam* null background. Transgene was recombined onto a $bam^{\Delta 86}$ chromosome by McKearin. Construction data can be found in (McKearin and Spradling, 1990)

the results were consistent across different insertion sites and in different *bam* genetic backgrounds (Figure 3.4A,C). We also determined that *bam* expression in the stock from which the *bam* allele in *mel-bam-yfp* was cloned is similar to the *D. melanogaster bam* heterozygote, demonstrating that the particular allele we chose is not defective in expression (Figure 3.4B). Additionally, we found that *bam* expression in the heterozygous genotype used as a reference is not an outlier as it is similar across several genetic backgrounds (Figure 3.4B). Finally, we compared *bam* expression in *mel-bam-yfp;bam*$^{-}$ to that of another *bam* transgene, *bam-α;bam*$^{-}$ (McKearin and Spradling, 1990) previously reported to fully complement both female and male sterility of *D. melanogaster bam* mutants. We found the *bam-α* transgene is similarly under-expressed relative to the *D. melanogaster bam* heterozygote (Figure 3.4A). Overall, these results demonstrate that neither our *mel-bam-yfp* nor other *bam* transgenes express at a wildtype level in females. In contrast to females, *bam* expression levels in *mel-bam-yfp;bam*$^{-}$ males are similar to the *D. melanogaster bam* male heterozygote. Additionally, while *bam* expression in *sim-bam-yfp;bam*$^{-}$ males is higher than *bam* levels in both *mel-bam-yfp;bam*$^{-}$ the *D. melanogaster bam* heterozygote, it is not significantly higher and will be discussed in further detail in section 3.3.3. The difference in expression levels between the transgenes complicates the ability to attribute phenotypic differences between the orthologs to coding sequence divergence. I therefore expand our analyses to include the *D. melanogaster bam* heterozygote as an expression control, since its expression level is comparable to *bam* levels in *sim-bam-yfp;bam*$^{-}$, resulting in a three-way comparison: *mel-bam-*

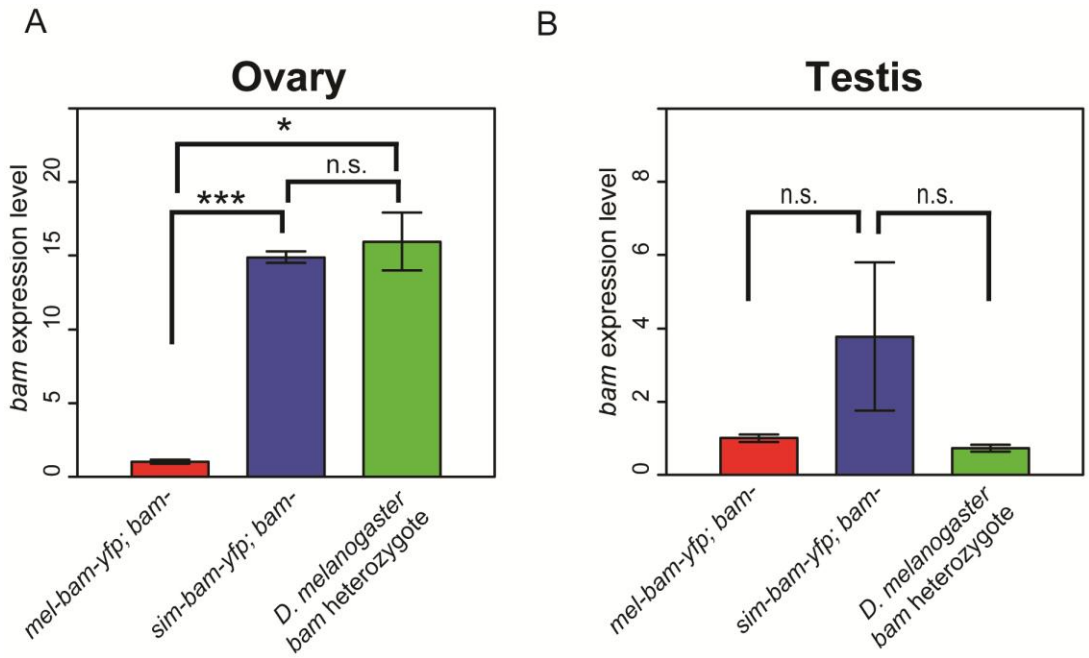
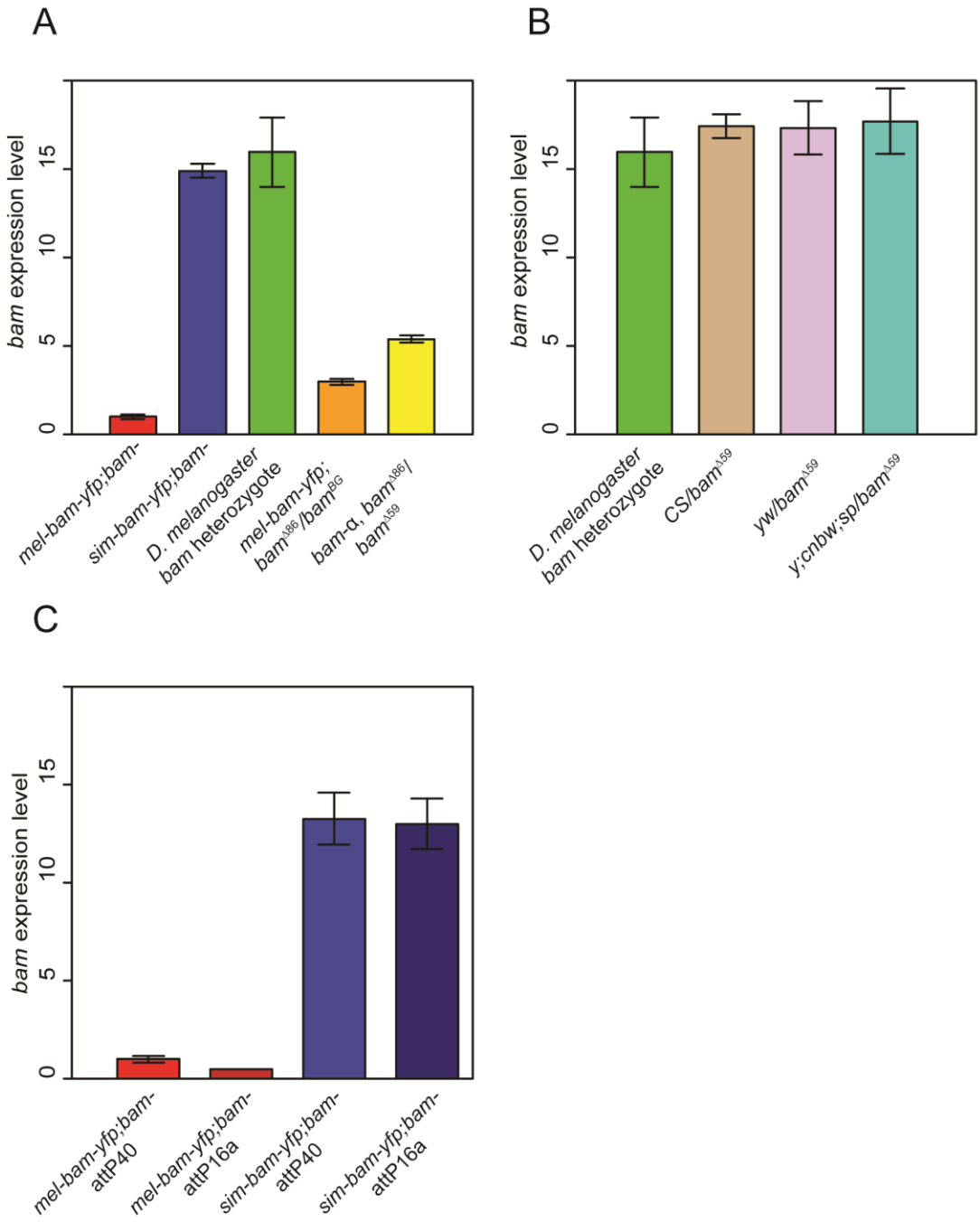


Figure 3.3: qRT-PCR of *bam* transgenes and the *D. melanogaster bam* heterozygote. (A) *mel-bam-yfp* is significantly underexpressed in ovaries. (B) *sim-bam-yfp* is not significantly different in expression from *mel-bam-yfp* or the *D. melanogaster bam* heterozygote in testes. *bam* RNA levels from ovaries or testes are shown. In transgenic *D. melanogaster* ovaries, *sim-bam-yfp*;*bam*$^{-}$ (blue) is expressed at a level equivalent to that of the *D. melanogaster bam* heterozygote (green), but *mel-bam-yfp*;*bam*$^{-}$ (red) is expressed at a much lower level. In transgenic *D. melanogaster* testes, *sim-bam-yfp*;*bam*$^{-}$ is more abundant than either *mel-bam-yfp*;*bam*$^{-}$ or the *D. melanogaster bam* heterozygote. Each transgene is present in 1 copy and should be equivalent in expression to the *D. melanogaster bam* heterozygote. Transgenes are inserted at attP40. N = 3 biological replicates for each sample. (T-test, *P < 0.025, **P < 0.01, ***P < 0.001).

Figure 3.4: qRT-PCR of *bam* expression controls.
(A) Underexpression of *bam* in *mel-bam-yfp; bam*$^{-}$ ovaries is not due to genetic background or the YFP tag. Ovarian *bam* RNA levels from *mel-bam-yfp* (red) and *sim-bam-yfp* (blue) in the *bam* mutant background (*bam*$^{\Delta 86}$/*bam*$^{\Delta 59}$), and the *D. melanogaster bam* heterozygote (green) from figure 3.3A are shown in addition to other genotypes. *bam* levels of *mel-bam-yfp* in a different *bam* mutant background (*bam*$^{\Delta 86}$/*bam*^{BG}) (orange) are also reduced relative to the *bam* heterozygote. *bam* RNA levels of a different *bam* transgene (yellow, *bam*-α;*bam*$^{-}$) is reduced in expression. ΦC31-integrated transgenes shown in (A) are docked in attP40.
(B) *bam* expression levels show little variation across strains. *bam* RNA levels were compared to the *D. melanogaster bam* heterozygote shown in (A) to that of various wildtype or marker lines (*Canton S*, *yw*, *y;cn bw;sp*) that were made heterozygous over a *D. melanogaster bam* mutant (*bam*$^{\Delta 59}$). The *bam* sequence in *mel-bam-yfp* was cloned from *y;cn bw;sp*. (C) Transgene expression is stable across different insertion sites. We compared *bam* RNA levels from *mel-bam-yfp;bam*$^{-}$ and *sim-bam-yfp;bam*$^{-}$ ovaries in two different insertion sites, attP40 and attP16a. attP40 samples are the same as those shown in part A. N = 3 biological replicates for each sample.


A
bam expression level
0
5
10
15
mel-bam-yfp;bam-
sim-bam-yfp;bam-
D. melanogaster bam heterozygote
mel-bam-yfp; bamΔ86/bamBG
bam-α, bamΔ86/ bamΔ59
B
bam expression level
0
5
10
15
D. melanogaster bam heterozygote
CS/bamΔ59
yw/bamΔ59
y;cnbw;sp/bamΔ59
C
bam expression level
0
5
10
15
mel-bam-yfp;bam- attP40
mel-bam-yfp;bam- attP16a
sim-bam-yfp;bam- attP40
sim-bam-yfp;bam- attP16a

yfp;bam^{-} vs. *bam* heterozygote, *mel-bam-yfp;bam*$^{-}$ vs. *sim-bam-yfp;bam*$^{-}$, and *sim-bam-yfp;bam*$^{-}$ vs *bam* heterozygote. In the discussion, I consider possible causes of these surprising results on transgene expression levels and how they impact our conclusions about the effects of sequence divergence on *bam* function.

3.3.2 *D. simulans bam* cannot fully complement *D. melanogaster bam*'s function in female fertility

To assay the ability of the transgenes to complement both female and male sterility, we crossed each transgene into a *D. melanogaster bam* transheterozygous, null mutant background. Sibling flies generated from this cross that were heterozygous for *bam* but did not carry a transgene were also used as a comparison in fertility experiments. We find that *mel-bam-yfp* fully rescues *D. melanogaster bam* female sterility to the level of the *D. melanogaster bam* heterozygous control. This result suggests that the transgene is fully functional in females despite having a reduced expression level relative to wildtype *bam* alleles. However, *sim-bam-yfp;bam*$^{-}$ females were significantly less fertile than either *mel-bam-yfp;bam*$^{-}$ or the *D. melanogaster bam* heterozygous control class throughout the duration of the experiment (Figure 3.5), demonstrating the *sim-bam-yfp* cannot fully rescue *D. melanogaster bam* female sterility.

3.3.3 *sim-bam-yfp* rescues male sterility

In contrast to female fertility assays, we found that *sim-bam-yfp;bam*$^{-}$ males were as fertile as their *mel-bam-yfp;bam*$^{-}$ or *D. melanogaster bam* heterozygous

counterparts (Figure 3.6A). To test for more subtle differences in male fertility, we used a sperm exhaustion mating assay by providing the males with new virgin females every day over a five day period. Under these more stringent conditions we found that *sim-bam-yfp* fully rescues male sterility compared to the *D. melanogaster* heterozygous control, but *mel-bam-yfp* does not (Figure 3.6B). This result was unexpected based on *bam* RNA levels from testes. As shown in Figure 3.3, *bam* expression levels in *mel-bam-yfp;bam*$^{-}$ are similar to both the *D. melanogaster bam* heterozygote and *sim-bam-yfp;bam*$^{-}$. We performed more qRT-PCR to ask if these expression patterns are consistent across different insertion sites. While *bam* expression in *sim-bam-yfp;bam*$^{-}$ at attP40 is extremely high and variable compared to the other genotypes, the overall trends are similar; *bam* expression in *mel-bam-yfp;bam*$^{-}$ is expressed at similar levels to the *D. melanogaster* heterozygote, while expression in *sim-bam-yfp;bam*$^{-}$ is higher than both, although not significantly so (Figure 3.7A). Furthermore, we also compared the transgenic *bam* expression to *bam* expression in the wildtype background into which they were crossed, *yw*, as well as to the *bam-α;bam*$^{-}$ transgene. We find that *D. melanogaster bam* expression in *yw* is roughly double that of the heterozygous control, but that *bam-α;bam*$^{-}$ is expressed much higher than both. This result is surprising because the *bam-α;bam*$^{-}$ transgene acts similarly to *mel-bam-yfp;bam*$^{-}$ in male sperm exhaustion experiments in that they become sterile very quickly (Figure 3.7B). These results suggest that *bam* RNA levels are much more variable in testes than in ovaries and may also suggest that there may be little correlation between *bam* expression level and activity in males. Alternatively,

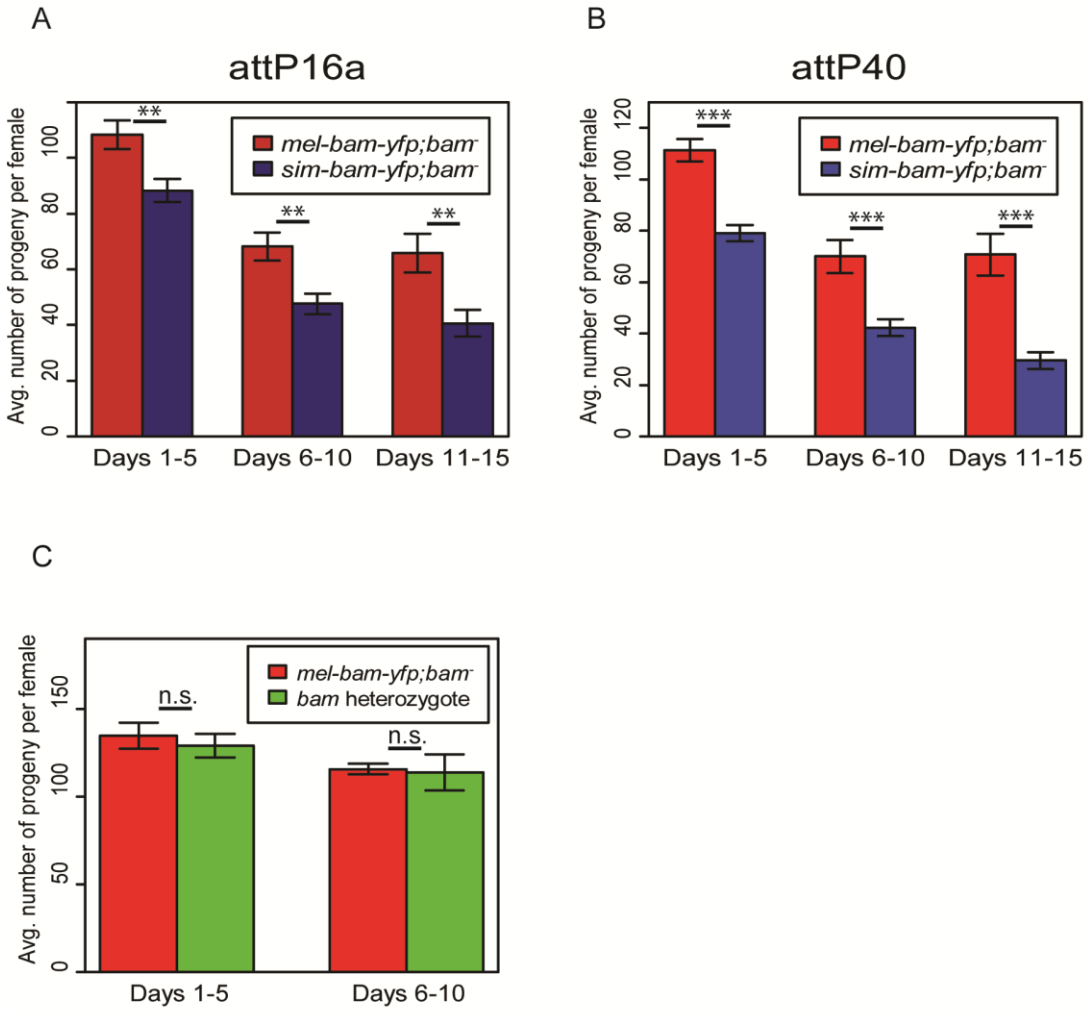


Figure 3.5: *sim-bam-yfp* does not fully rescue *D. melanogaster bam* mutant female sterility. One transgenic female (or heterozygous female) and two tester males were allowed to mate and the trio was transferred to a new vial every five days. Fertility is shown as the average number of progeny per female +/- SEM for each vial. (A-B) *sim-bam-yfp* cannot fully rescue *D. melanogaster bam* female sterility. *mel-bam-yfp;bam*$^{-}$ is shown in red and compared to *sim-bam-yfp;bam*$^{-}$ in blue. (A) Data generated from transgenes inserted in attP16a. (B) Data generated for transgenes inserted in attP40. (C) *mel-bam-yfp* rescues *D. melanogaster bam* sterility. *mel-bam-yfp;bam*$^{-}$ is shown in red and the *D. melanogaster bam* heterozygote is shown in green. (T-test, * P<0.05, **P<0.01, ***P<0.001).

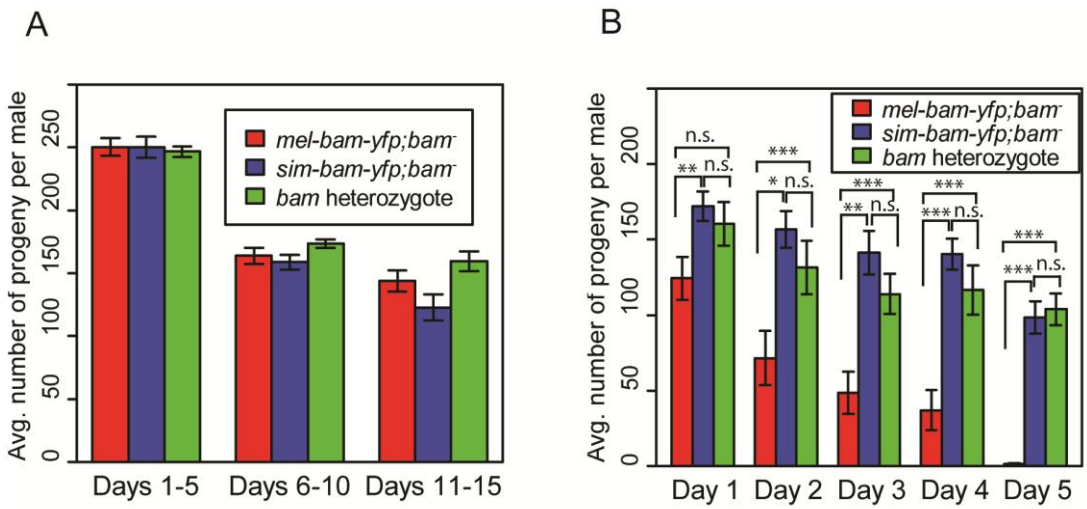


Figure 3.6: *sim-bam-yfp* rescues *D. melanogaster bam*⁻ male sterility.
(A) *mel-bam-yfp* and *sim-bam-yfp* both rescue male sterility under standard fertility conditions. One male and two tester females were allowed to mate and the trio was transferred to a new vial every five days. No comparisons are significantly different in (A). (B) *sim-bam-yfp* but not *mel-bam-yfp* rescues male sterility under sperm exhaustion conditions. One male was allowed to mate with a fresh pair of virgin tester females everyday for five days total. Male fertility is shown as the average number of progeny per male +/- SEM for each vial. *mel-bam-yfp;bam*⁻ is shown in red, *sim-bam-yfp;bam*⁻ in blue, and the *D. melanogaster bam* heterozygote is shown in green. Transgenes are inserted in attP40. (T-test, *P<0.05, **P<0.01, ***P<0.001).

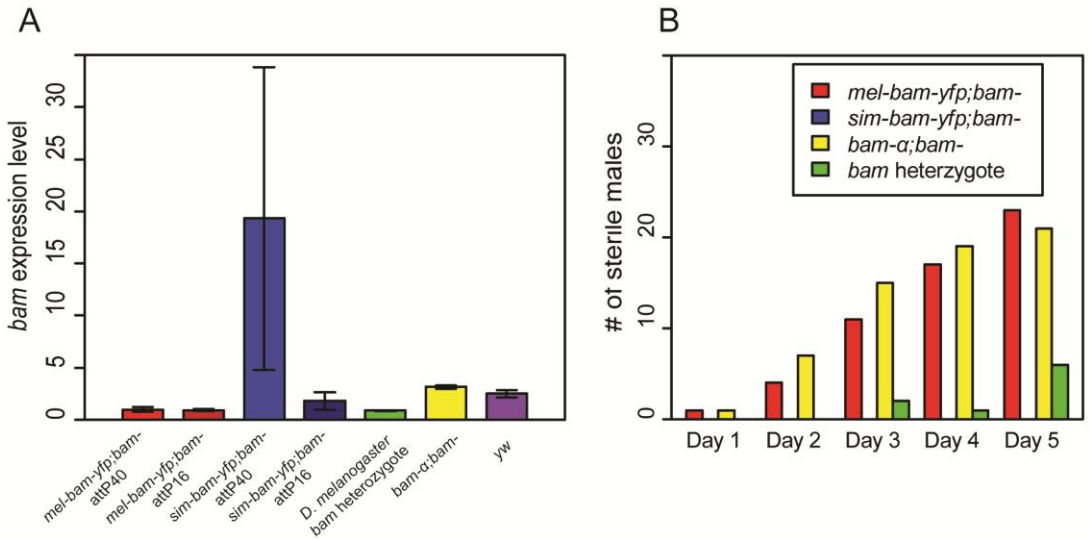


Figure 3.7: *bam* RNA levels in testes show inconsistent results and do not correlate with male fertility assays. (A) *bam* RNA levels from *mel-bam-yfp;bam*$^{-}$ (red) and *sim-bam-yfp;bam*$^{-}$ (blue) from two different insertion sites. The level of *bam* expression in the *D. melanogaster bam* heterozygote (green) is also shown. *bam* RNA levels of a different *bam* transgene present in single copy is also shown (yellow, *bam*-α;*bam*$^{-}$). N = 3 biological replicates. (B) Number of sterile males during sperm exhaustion experiment. Number of sterile males from *mel-bam-yfp;bam*$^{-}$ (red), *sim-bam-yfp;bam*$^{-}$ (blue), *D. melanogaster bam* heterozygote (green), and *bam*-α;*bam*$^{-}$(yellow). Transgenes (not including *bam*-α) are in attP40. While *bam* expression in *mel-bam-yfp;bam*$^{-}$ is the same as the *D. melanogaster bam* heterozygote, this transgene fails to rescue male sterility. Similarly, *bam* expression in *bam-α;bam*$^{-}$ is higher than the *D. melanogaster bam* heterozygote and similar to the wildtype line, *yw*, which has two copies of *bam*, yet it does not fully rescue *bam* male sterility.

the qRT-PCR threshold values (Ct) of *bam* testes samples were higher than generally suggested (28-32) and could not produce a standard curve with good efficiency, so the samples were quantified based on a standard curve generated from ovary cDNA. Furthermore, -RT controls had Ct values that were only 1-2 cycles higher than +RT samples. At higher Ct values, estimation of RNA levels may not be as reliable, so these results may reflect experimental error rather than meaningful biology. Nonetheless, our transgenic assay system has shown evidence for functional divergence between the *D. melanogaster* and *D. simulans bam* orthologs and that this divergence predominantly affects *bam*'s role in female fertility.

3.3.4 Ovaries from *sim-bam-yfp;bam*$^{-}$ females show multiple defects including GSC loss

To determine the cause of the reduced fertility of *sim-bam-yfp;bam*$^{-}$ females, we stained *mel-bam-yfp;bam*$^{-}$ and *sim-bam-yfp;bam*$^{-}$ ovaries with antibodies to the germline marker Vasa, the fusome marker Hts-1B1, and to the YFP tag in Bam-YFP. The ovaries of flies with *mel-bam-yfp;bam*$^{-}$ show wildtype morphology (Figure 3.8A,B). GSCs were identified by their spherical fusome (i.e. the spectrosome) and their location within the germarium. *mel-bam-yfp;bam*$^{-}$ovaries had 2-3 GSCs per germarium, which is comparable to wildtype levels, and Bam was properly localized (McKearin and Ohlstein, 1995; Kirilly and Xie, 2007). Furthermore, the vast majority of egg chambers underwent the proper number of cyst divisions giving rise to 16-cell cysts (Table 3.3). In contrast, ovaries from *sim-bam-yfp;bam*$^{-}$ flies showed multiple ovarian defects that increased as the flies aged (Figure 3.8C,D). First, they exhibit

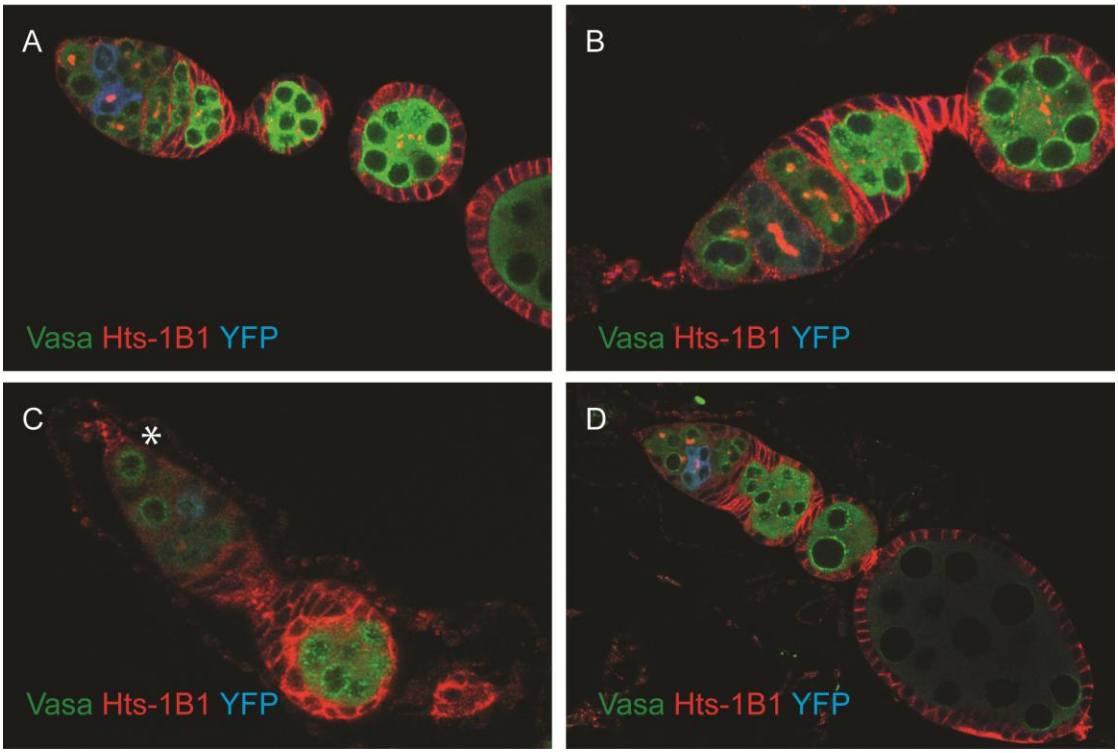


Figure 3.8: *sim-bam-yfp;bam*$^{-}$ ovaries show multiple defects.
(A-B) *mel-bam-yfp;bam*$^{-}$ ovaries show wildtype morphology including proper Bam-YFP expression, correct number of GSCs marked by spherical fusomes, and proper numbers of cells/cyst. (C-D) *sim-bam-yfp;bam*$^{-}$ ovaries show reduced number of GSCs (*) and contain egg chambers with improper number of cells/cyst. (A-D) Ovaries are stained with antibodies to Vasa (green), Hts-1B1 (red), and YFP (blue).

stem cell loss, with an average of 1.5 GSCs per ovariole when young (days 1-5, Table 3.3) that appeared to decrease with age, although this was not quantified. Second, as the flies age (days 6-15) they have a reduction in the number of ovarioles containing mature egg chambers as a consequence of GSC loss. Also, many of the egg chambers that are present have an improper number of cyst divisions and show mitotic synchrony defects (Table 3.3). Mitotic synchrony defects are typically seen with other fusome mutants (e.g. hts (Yue and Spradling, 1992) and α-spectrin (de Cuevas et al., 1996)) suggesting that *sim-bam-yfp;bam*$^{-}$ flies may have fusome defects. However, *sim-bam-yfp;bam*$^{-}$ ovaries have both reduced and increased numbers of cyst divisions while other fusome mutants have only reduced numbers, suggesting instead that *sim-bam-yfp* cannot properly regulate the number of cyst divisions, independently of potential fusome defects. Despite these multiple ovarian defects, it is important to note that sim-Bam-YFP shows a proper localization pattern. It is absent in GSCs and present in mitotically active cysts, suggesting that the defects are not due to gross misregulation of Bam. Moreover, *sim-bam-yfp;bam*$^{-}$ flies never show the *D. melanogaster bam* null mutant phenotype of tumorous ovaries (McKearin and Spradling, 1990), suggesting that *sim-bam-yfp* is capable of rescuing the GSC differentiation defect in *D. melanogaster bam* mutant females.

Table 3.3: *sim-bam-yfp;bam*$^{-}$ flies have multiple ovarian defects

Transgene	Transgene Dose	Avg. # of GSCs	# of egg chambers with indicated # of cells/cyst		
			<16	16	>16
mel-bam-yfp;bam^{-}	1	2.8	2	97	1
D. melanogaster bam heterozygote	0	2.9	1	96	3
sim-bam-yfp;bam^{-}	1	1.5	10	76	14

Ovaries were dissected from flies aged for 3-5 days post-eclosion on yeast. N = at least 50 ovarioles for each sample.

3.3.5 *D. simulans* Bam properly interacts with *D. melanogaster* Bgcn

Bam is adaptively evolving at a high rate in both *D. melanogaster* and *D. simulans,* having more than four times the number of amino acid substitutions than the average seen between these two species (Civetta et al., 2006; Bauer DuMont et al., 2007). In *D. melanogaster*, Bam and Bgcn physically interact (Li et al., 2009; Shen et al., 2009), and *bgcn* is also evolving under rapid, adaptive evolution in both *D. melanogaster* and *D. simulans* (Bauer DuMont et al., 2007). With both proteins adaptively evolving in both species, one would expect that if substitutions occurred that reduce their interaction, compensatory mutations would be selected for to re-establish a strong interaction. Therefore, independent substitutions and compensatory substitutions occurring at Bam and Bgcn within each species might render the protein partners incapable of, or less efficient at, interacting when brought together with the heterospecific protein. To determine if sim-Bam and mel-Bgcn are capable of interacting with one another, we performed immunoprecipitation assays from Drosophila S2 cells. Cells were transiently transfected with either mel-Bam::HA or sim-Bam::HA, and with mel-Bgcn::MYC transgenes. We found that in reciprocal immunoprecipitation experiments both the conspecific and heterospecific Bam coimmunoprecipatated with mel-Bgcn::MYC, indicating that sim-Bam can interact with mel-Bgcn (Figure 3.9).

These assays involve gene over-expression and cannot discriminate whether the interaction is reduced in efficacy. Ohlstein et al. (2000) showed that *bgcn* acts as a dominant enhancer of *D. melanogaster bam* phenotypes. In this assay *bam*

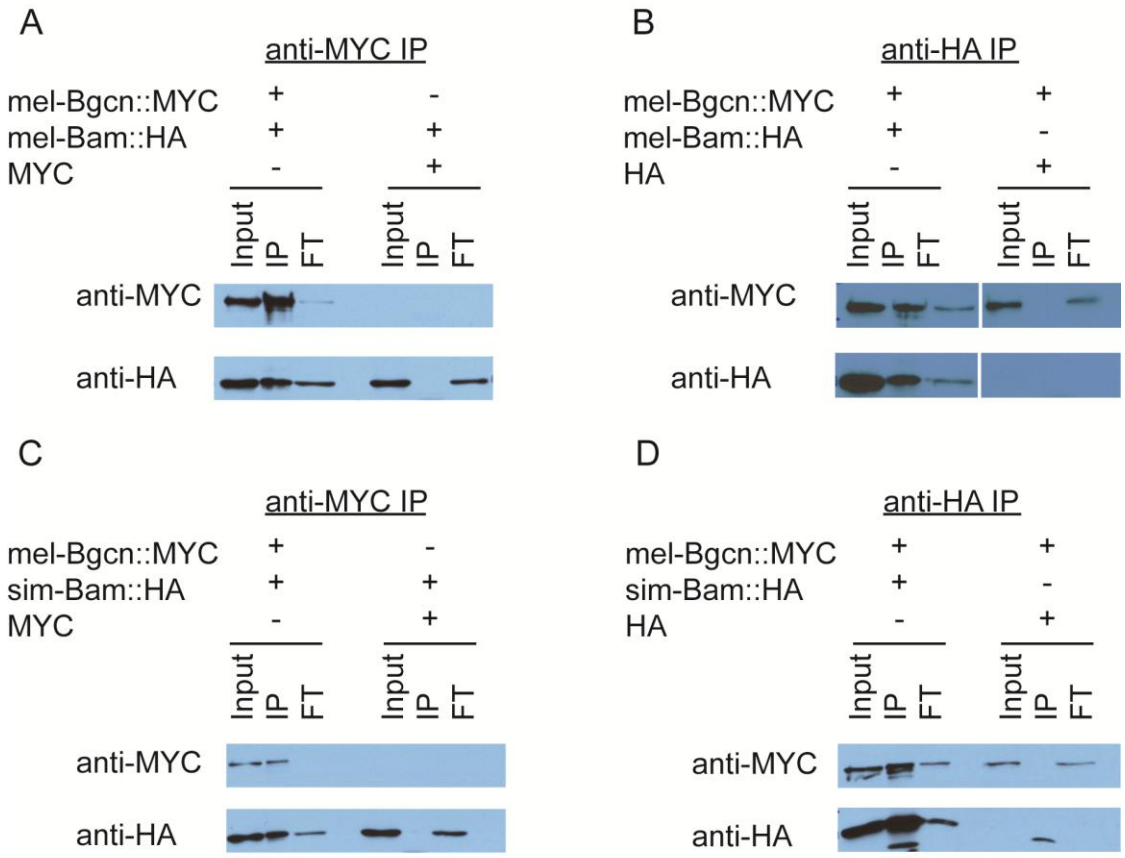


Figure 3.9: sim-Bam can reciprocally immunoprecipitate with mel-Bgcn. (A) Control experiment showing that mel-Bam::HA immunoprecipitates with mel-Bgcn::MYC but not with MYC alone. Cells were transfected with mel-Bam::HA and either mel-Bgcn::MYC or MYC. Anti-MYC immunoprecipitates from Drosophila S2 cells were analyzed by western blot. (B) Control experiment showing that mel-Bgcn::MYC immunoprecipitates from mel-Bam::HA but not HA alone. Cells were transfected with mel-Bgcn::MYC and either mel-Bam::HA or HA. Anti-HA immunoprecipitates were analyzed by western blot. (C) sim-Bam::HA immunoprecipitates with mel-Bgcn::MYC but not with MYC alone. Cells were transfected with sim-Bam::HA and either mel-Bgcn::MYC or MYC. Anti-MYC immunoprecipitates were analyzed by western blot. D) sim-Bgcn immunoprecipitates from mel-Bam::HA but not HA alone. Cells were transfected with mel-Bgcn::MYC and either sim-Bam::HA or HA. Anti-HA immunoprecipitates were analyzed by western blot. Smaller bands found in IP lanes when probed with anti-HA correspond to antibody heavy chains. Gels are loaded with 25% of total input, 100% of immunoprecipitate (IP), and 10% of protein that did not immunoprecipitate (flow through, FT) was loaded in each respective lane.

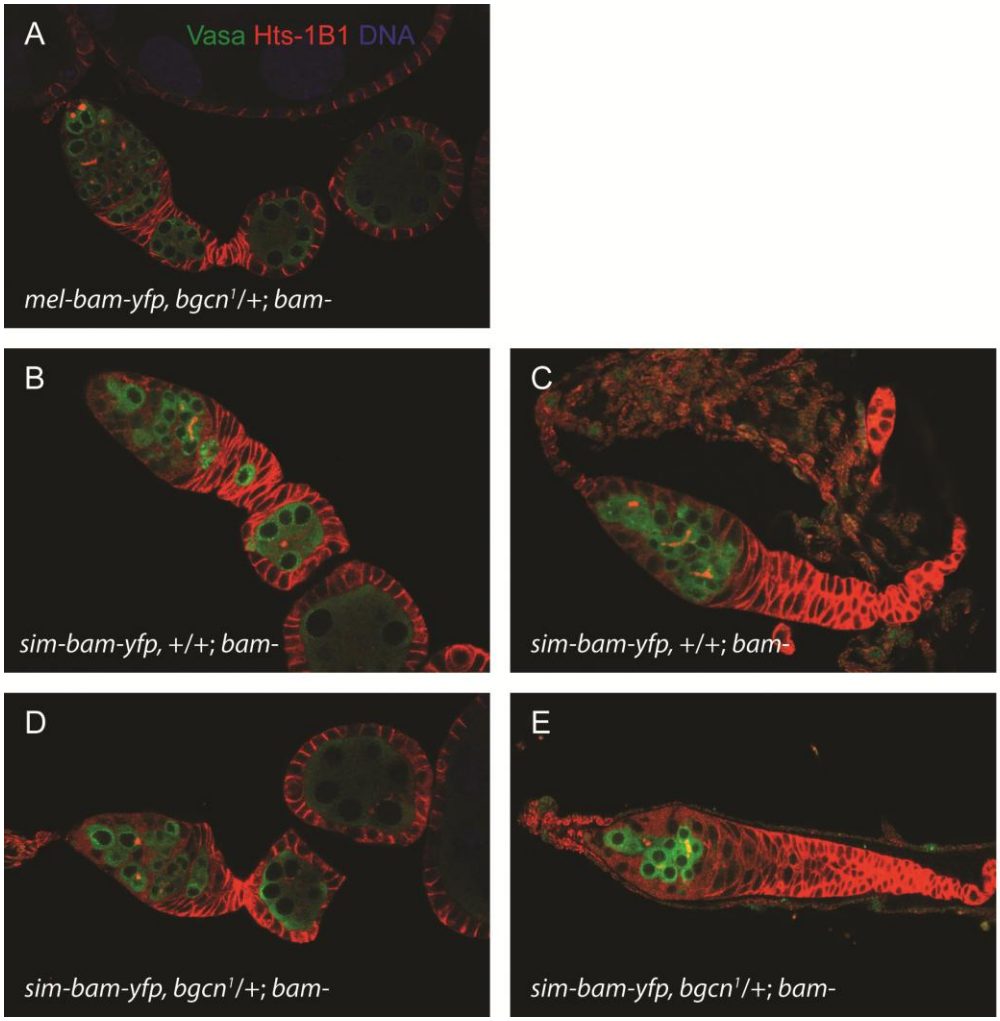


Figure 3.10: Reduction of *bgcn* dosage does not exacerbate *sim-bam-yfp;bam*$^{-}$ ovarian defects. (A) *mel-bam-yfp;bam*$^{-}$ flies are insensitive to changes in *bgcn* levels, as ovaries have wildtype morphology and no GSC tumors. This is consistent with previously reported data that flies doubly heterozygous for *bam* and *bgcn*, show wildtype fertility (Ohlstein et al., 2000). (B-C) Ovaries of *sim-bam-yfp;bam*$^{-}$ flies show a varying range of ovarian defects with mild (B) and moderate (C) examples shown. (D-E) Removal of a copy of *bgcn* (*bgcn*^{1}) does not enhance the range of phenotypes seen in *sim-bam-yfp;bam*$^{-}$ ovaries. No tumorous ovaries were ever seen (N > 50 ovarioles) suggesting that *sim-bam-yfp* and *D. melanogaster bgcn* function normally together in promoting GSC differentiation. (A-E) Ovaries are stained with antibodies to Vasa (green), Hts-1B1 (red), and with DAPI (Blue).

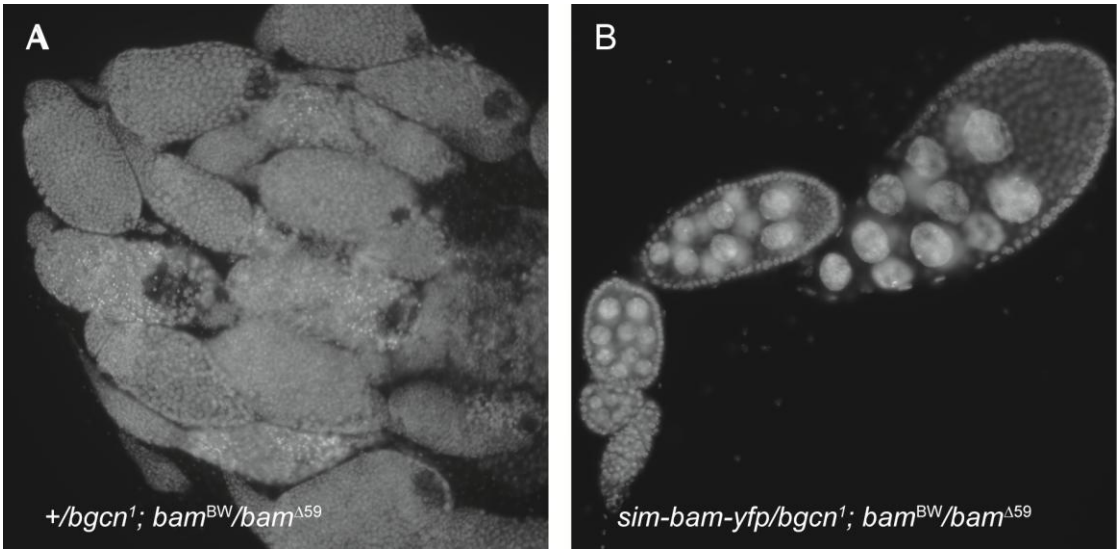


Figure 3.11: *sim-bam-yfp* rescues the tumorous phenotype seen in the *bam* hypomorph-*bgcn* interaction. (A) As described in Ohlstein et al. (2000), removal of one copy of *bgcn* exacerbates a *bam* hypomorph resulting in completely tumorous ovaries. The egg chambers of these ovaries are filled with small nuclei. (B) The addition of one copy of *sim-bam-yfp* suppresses the tumorous ovary defects, suggesting that *sim-bam-yfp* can properly interact with *D. melanogaster bgcn*. Ovaries are stained with DAPI.

hypomorphic mutations resulting in weakly fertile females were used. It was found that reducing *bgcn* dosage exacerbated the *bam* phenotype, causing sterility and giving rise to completely tumorous ovaries. We reduced the copy number of *bgcn* by half (*bgcn*^{1}/+) and found no exacerbation of the *sim-bam-yfp* phenotype (Figure 3.10D,E). Additionally, we complemented the *bgcn*-induced sterility of a *bam* hypomorph by adding a copy of *sim-bam-yfp* (Figure 3.11). Together the co-immunoprecipitation and genetic-interaction experiments suggest that *sim-bam-yfp;bam*$^{-}$ ovarian defects are not due to an inability of sim-Bam to interact with mel-Bgcn.

3.3.6 *sim-bam-yfp* ovarian defects are dose dependent

bam mutants give rise to tumorous ovaries due to the inability of GSCs to differentiate. In contrast, *sim-bam-yfp* appears to act as a gain-of-function allele by causing GSC loss, suggesting that mis-expression or mis-regulation of *sim-bam-yfp* levels causes the ovarian defects. We did not observe sim-Bam-YFP mislocalization in the above experiments, but IHC may not be sensitive enough to identify small changes in expression. Therefore, we asked if adding additional copies of the *mel-bam-yfp* or *sim-bam-yfp* transgenes either improves or worsens the fertility phenotypes. We found no significant differences in fertility when comparing *mel-bam-yfp;bam*$^{-}$ to 2x *mel-bam-yfp;bam*$^{-}$. However, 2x *sim-bam-yfp;bam*$^{-}$ transgenic flies showed a significant decrease in fertility when compared to *sim-bam-yfp;bam*$^{-}$ and were nearly sterile by day 15 (Figure 3.12A,B)

Ovarioles from 2x *sim-bam-yfp;bam*$^{-}$ flies showed accelerated rates of stem cells loss, even in young (1-5 day old) flies (Figure 3.12, Table 3.4). 2x *sim-bam-*

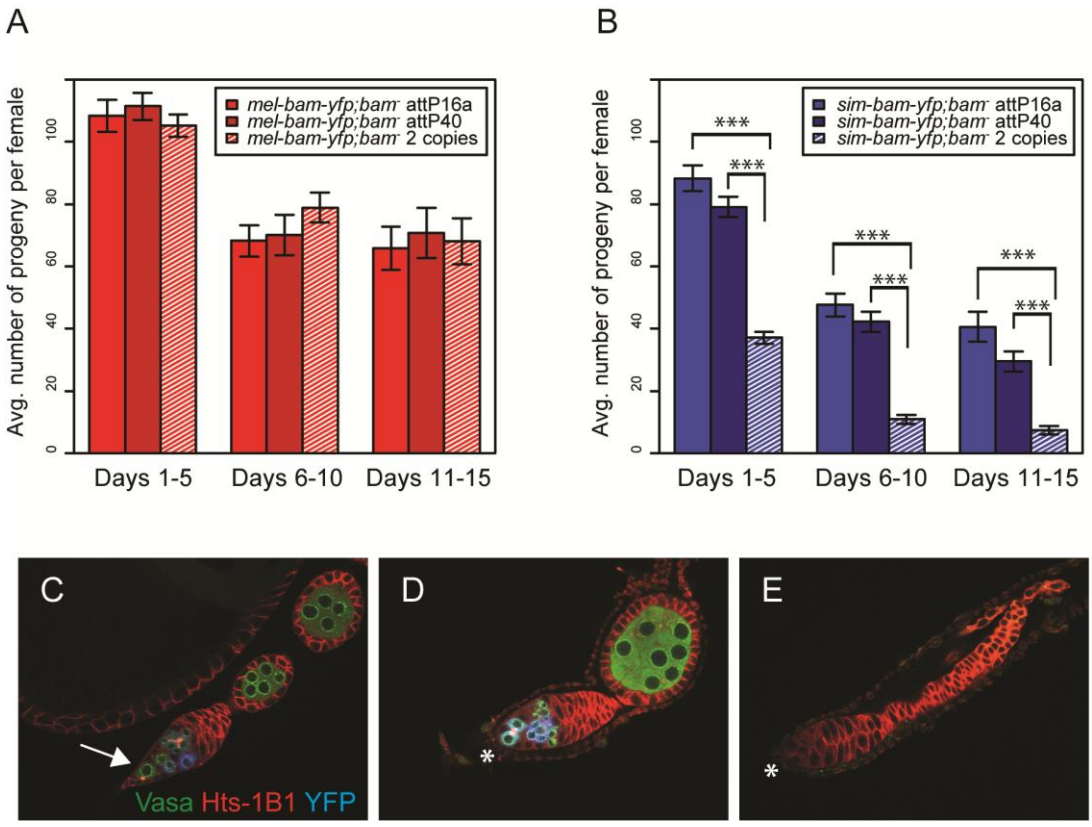


Figure 3.12: *sim-bam-yfp:bam*$^{-}$ female fertility decreases with additional copies of *sim-bam-yfp*. (A*) mel-bam-yfp;bam*$^{-}$ is not dose sensitive (no comparison is stastically significantly different). (B) *sim-bam-yfp;bam*$^{-}$ is dose sensitive. For A and B, one female and two tester males were allowed to mate and the trio was transferred to a new vial every five days. Fertility is shown as the average number of progeny per female +/- SEM for each vial. All *mel-bam-yfp;bam*$^{-}$ comparisons are not significant while *sim-bam-yfp;bam*$^{-}$ comparisons are highly significant (T-test, ***P<0.001). (C) Ovaries of flies with *sim-bam-yfp;bam*$^{-}$ show evidence of GSC loss but still contain GSCs. Arrow indicates single GSC. (D) Ovaries of 2x *sim-bam-yfp;bam*$^{-}$ flies show increased GSC loss, as most ovarioles no longer contain GSCs but still have germline present (Vasa staining) and (E) some have completely lost all GSCs and germline. Asterisk indicates anterior end of germarium where GSCs would be. For C-E, flies are stained for Vasa (green), Hts-1B1 (red), and YFP (blue).

Table 3.4: GSC number in transgenic lines

Transgene	# of GSCs
2x *mel-bam-yfp;bam*$^{-}$	2.88
sim-bam-yfp;bam/+	2.60
2x *sim-bam-yfp;bam*$^{-}$	0.5
2x *sim-bam-yfp;bam*/+	2.33

yfp;bam^{-} flies typically lacked GSCs and in some cases they no longer contained any germline cells, as seen by lack of Vasa staining (Figure 3.12), while in *sim-bam-yfp;bam*$^{-}$ flies, GSCs were almost always present in every ovariole though often reduced in number.

We performed qRT-PCR comparing the ovarian RNA expression levels of *bam* from flies with one or two copies of the transgene. As reported above, *bam* levels in *sim-bam-yfp;bam*$^{-}$ increase approximately 13-fold relative to *bam* levels in *mel-bam-yfp;bam*$^{-}$ (Figure 3.13). As expected, doubling the dose of the transgenes results in a doubling of expression for both *mel-bam-yfp* and *sim-bam-yfp* (Figure 3.13). Notably, however, *bam* RNA levels of 2x *sim-bam-yfp;bam*$^{-}$ are not greater than in *D. melanogaster* wildtype flies (Figure 3.13). We conclude that the 2x *sim-bam-yfp;bam*$^{-}$ defects are specifically due to increased dosage of the functionally diverged *D. simulans bam*, rather than to a general effect of increasing *bam* dosage.

3.3.7 sim-Bam-YFP Protein is slightly more abundant than mel-Bam-YFP

To further investigate the cause of the deleterious effects of *sim-bam-yfp;bam*$^{-}$, we compared Bam-YFP protein levels from ovarian lysates of *mel-bam-yfp;bam*$^{-}$ and *sim-bam-yfp;bam*$^{-}$. Using an antibody to YFP, we find that flies with *sim-bam-yfp* have slightly more Bam-YFP protein than those with *mel-bam-yfp* (~1.2 fold difference, Figure 3.14), but this difference is much smaller than what we see at the level of RNA expression (~13 fold difference, Figure 3.3). This result supports our inference above that the deleterious effects of seen in *sim-bam-yfp;bam*$^{-}$ flies are due to divergence of *D. simulans* Bam protein and not to overexpression. The recovery in

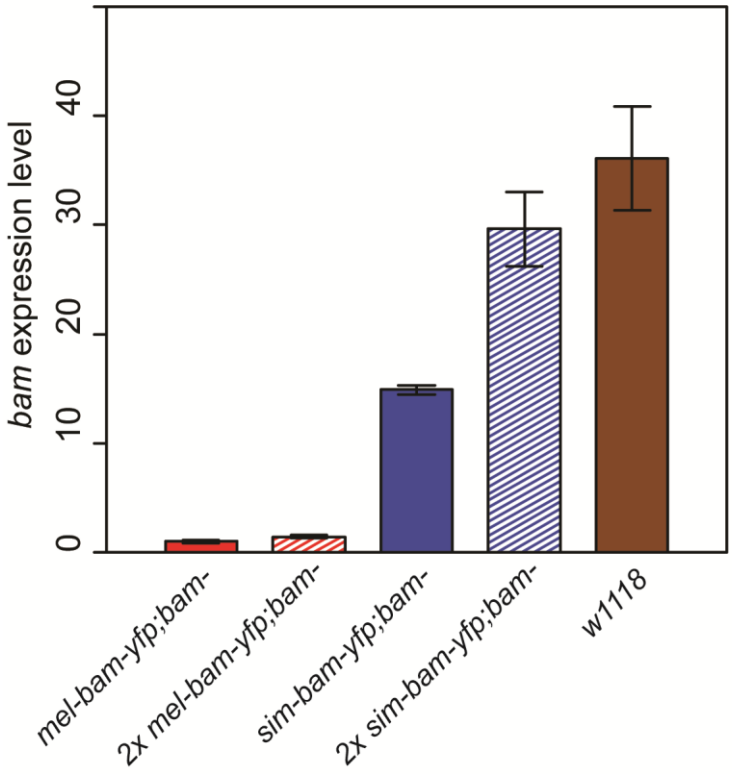


Figure 3.13: qRT-PCR of 1 and 2 copies of *mel-bam-yfp* and *sim-bam-yfp*. Ovarian *bam* RNA levels from *mel-bam-yfp;bam-* (red), 2x *mel-bam-yfp;bam*$^{-}$ (red striped), *sim-bam-yfp;bam*$^{-}$ (blue), 2x *sim-bam-yfp;bam*$^{-}$ (blue striped) and w^{1118} (brown) are shown. The expression of each transgene doubles with the addition of a second transgene copy. N = 3 biological replicates for each genotype.

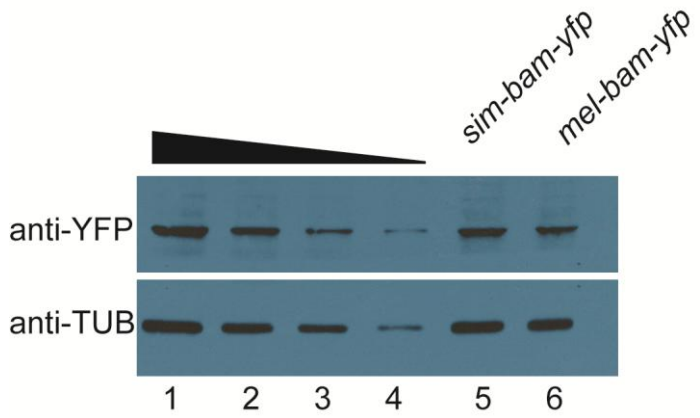


Figure 3.14: Semi-quantitative Western Blot comparing sim-Bam-YFP and mel-Bam-YFP levels. Western blot probed with antibodies against anti-YFP or anti-α-Tubulin. Lanes 1-4: Dilution series (20µg, 10µg, 5µg, 2.5µg) made from sim-Bam-YFP ovarian lysates. Lanes 5-6: 10µg of ovarian lysates from *sim-bam-yfp;bam*$^{-}$ (Lane 5) or *mel-bam-yfp;bam*$^{-}$ (Land 6) were compared. Quantification of each sample was made from the standard curve, as described in Shibata et al. (2011), resulting in an estimate of sim-Bam-YFP being 1.2 fold more abundant than mel-Bam-YFP. This experiment was performed one time. Afterward it was found that a single *bam* mutant stock was contaminated. I do not know if contamination may have affected these results. However, the differences seen in this experiment were similar to other Western Blots performed comparing YFP expression to Tubulin expression on these genotypes (2-3 fold), though these experiments did not include a standard curve.

the level of mel-Bam-YFP protein likely explains why we do not see fertility defects in *mel-bam-yfp;bam*$^{-}$ flies despite the low amount of *bam* RNA present.

3.3.8 Ovarian defects caused by *sim-bam-yfp* are not fully dominant

We have shown that *sim-bam-yfp;bam*$^{-}$ defects are dose-dependent and that these flies produce slightly more Bam-YFP protein than *mel-bam-yfp;bam*$^{-}$ flies. Our experiments above suggest that these effects of *sim-bam-yfp* are due to *D. simulans* divergence and not the increase in protein abundance. We further explored the nature of *sim-bam-yfp*-mediated defects by asking how they are modulated in the presence of a wildtype *D. melanogaster bam* allele. We envisioned 3 possible outcomes. The first is that if the effects are purely due to increased dosage then they should become worse with the addition of wildtype *D. melanogaster bam*. The second is that if they are purely neomorphic as a consequence of *D. simulans bam* divergence, then they should be unchanged. In other words *sim-bam-yfp* will be dominant over *D. melanogaster bam*. And the third is that if they are due to a failure of *bam* function due to divergence, then they should be complemented by *D. melanogaster bam* and thus be recessive.

We assayed our transgenes with the addition of endogenous copies of *D. melanogaster bam*. We found that *sim-bam-yfp;bam*$^{-}$-dependent defects are mostly alleviated by the addition of even a single endogenous copy of *D. melanogaster bam* (Figure 3.15). Flies with 2x *sim-bam-yfp;bam*$^{-}$ show severe fertility defects and increased GSC loss (See Figure 3.12D,E). Yet, the ovaries of these flies are restored to nearly wildtype morphology by the addition of a single endogenous copy of *D.*

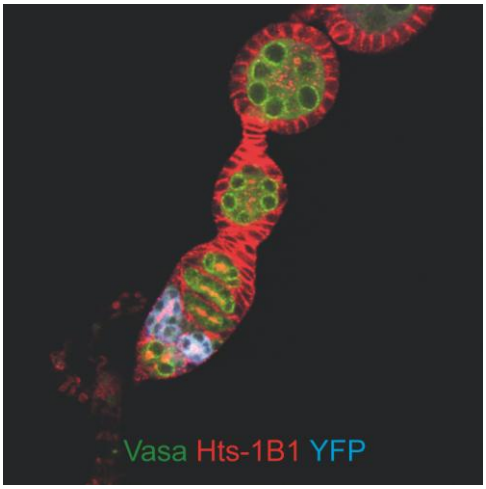


Figure 3.15: *D. melanogaster bam* rescues GSC loss of 2x *sim-bam-yfp;bam*$^{-}$ flies. Siblings of 2x *sim-bam-yfp;bam*$^{-}$ flies that contained one wildtype *D. melanogaster bam* allele (2x *sim-bam-yfp;bam/+*) were obtained and their ovaries are stained with antibodies to Vasa (green), Hts-1B1 (red), and YFP (blue). The ovaries of these flies show a much more wildtype morphology than 2x *sim-bam-yfp;bam*$^{-}$ (compare to 3.12 D, E for examples of ovaries of flies with 2x *sim-bam-yfp;bam*$^{-}$). Transgene is in attP40.

melanogaster bam (Figure 3.15, Table 3.4). These results suggest that *D. melanogaster bam* is dominant over *sim-bam-yfp*, but it is unlikely that *sim-bam-yfp* is simply a loss-of-function allele. First, its phenotypes do not match those seen in *bam* loss-of-function alleles in *D. melanogaster*. Second, the phenotypes in *sim-bam-yfp;bam*$^{-}$ ovaries become more severe as its dosage is increased. Third, the presence of *D. melanogaster bam* does not fully rescue *sim-bam-yfp* defects, suggesting that it may have both loss and gain of function attributes.

3.3.9 4x *sim-bam-yfp*;+/+ ovaries show ectopic Bam protein accumulation

An even more striking result came when looking at flies that contained four copies of *sim-bam-yfp* in a *D. melanogaster* fly wildtype for *bam* (4x *sim-bam-yfp*;+/+). In this background we observe some GSC loss, but much less than in 2x *sim-bam-yfp;bam*$^{-}$ flies. Furthermore, in a 4x *sim-bam-yfp;*+/+, we find ectopic sim-Bam-YFP protein that inappropriately accumulates in egg chambers that have budded off of the germarium (Figure 3.16A,B). In contrast, 4x *mel-bam-yfp*;+/+ flies show no evidence of ectopic Bam accumulation, although we did not quantitate protein levels in these 4x genotypes. One possible explanation for the ectopic sim-Bam-YFP protein accumulation is that *D. melanogaster* flies cannot properly recognize and degrade this protein. Alternatively, ectopic sim-Bam-YFP protein accumulation could occur if *D. melanogaster* flies simply have too much total Bam protein. In an effort to distinguish between these two alternatives, we co-stained the ovaries with antibodies to YFP and Bam. The Bam antibody only recognizes *D. melanogaster* Bam and not *D. simulans* Bam (Figure 3.17). If *D. melanogaster* flies simply have too much total Bam protein

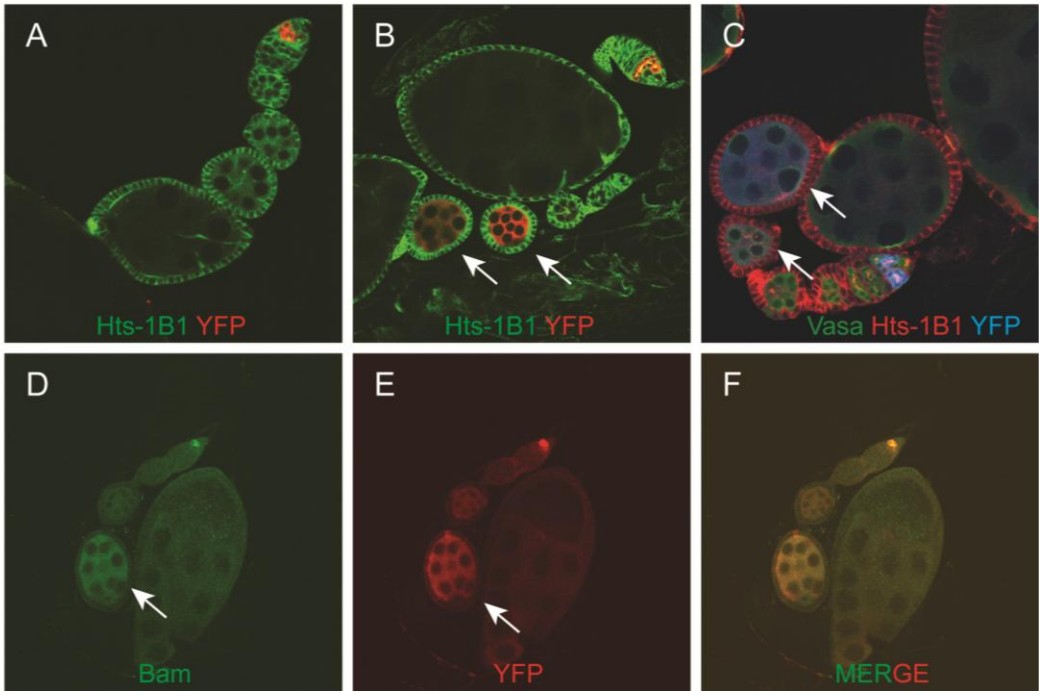


Figure 3.16: Ovaries of *D. melanogaster bam*+ flies with four copies of *bam-yfp* transgenes. (A) 4x *mel-bam-yfp:*+/+ show no ovarian defects, but (B) *4x sim-bam-yfp;*+/+ show ectopic sim-Bam-YFP localization in 16-cell cysts and egg chambers. (A-B) Ovaries are stained with antibodies to Hts-1B1 (green), and YFP (red). (C) *D. melanogaster bam*+ flies with four copies of *sim-bam-melpest-yfp* have ectopic Bam-YFP accumulation. Ovaries are stained with antibodies to Vasa (green), Hts-1B1 (red), and YFP (blue). (D-E) Ovaries from 4x *sim-bam-yfp;*+/+ flies. Ovaries are stained with antibodies to Bam (green) and YFP (red). (D-E) is an example of colocalization of Bam and YFP in egg chambers. White arrows point to ectopic Bam or Bam-YFP expression in egg chambers.

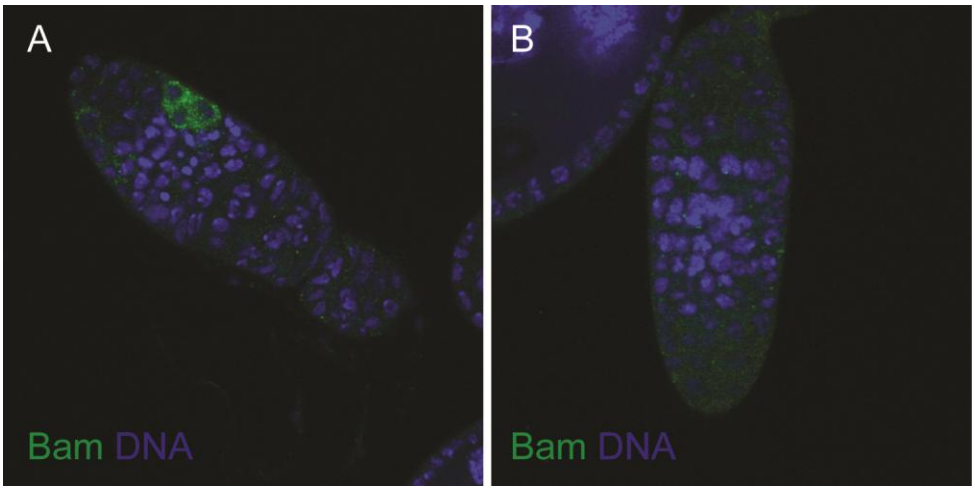


Figure 3.17: Monoclonal anti-Bam recognizes *D. melanogaster* Bam but not *D. simulans* Bam protein. (A) *D. melanogaster* ovaries. (B) *D. simulans* ovaries. Ovaries are stained with monoclonal anti-Bam (green) and DAPI (blue).

then one expects that both YFP and the Bam antibodies should stain egg chambers. However, if *D. melanogaster* flies are specifically unable to properly degrade sim-Bam-YFP protein, than only YFP should be present in egg chambers. We find that in many cases Bam and YFP antibodies colocalize in egg chambers (Figure 3.17C-E). Since sim-Bam-YFP is more abundant than mel-Bam-YFP (Figure 3.14), it appears that having four copies of specifically *sim-bam-yfp* in addition to two endogenous copies results in too much total Bam protein in the ovary that cannot be degraded efficiently.

3.3.10 sim-Bam-YFP ectopic expression is not due to PEST domain divergence

While 4x *sim-bam-yfp*;+/+ flies have ectopic Bam protein because they are inundated with too much total Bam protein, we do not see this for 4x *mel-bam-yfp*;+/+ flies, suggesting that small differences between Bam expression between these transgenes result in large protein differences when multiple transgenes are present. We were interested in testing the hypothesis that *D. melanogaster* might not be able to efficiently degrade sim-Bam-YFP protein, thus allowing for it to accumulate more than mel-Bam-YFP. While we cannot yet fully eliminate the possibility that sim-Bam-YFP protein is expressed at wildtype levels, our dose studies provided impetus to ask whether sequences in sim-Bam-YFP could cause it to not be degraded as efficiently. Bam contains a predicted PEST sequence (McKearin and Spradling, 1990) which is a protein motif thought to target proteins for rapid degradation (Rogers et al., 1986), and is commonly added to the C-terminus of proteins to increase their rate of turnover (Li et al., 1998; Frand et al., 2005). The 28-amino-acid-long PEST domain is important in Bam's regulation in spermatogenesis, as PEST domain deletions result in

increased Bam accumulation and fewer cyst divisions (Insco et al., 2009). The PEST domain of *D. melanogaster* Bam differs from *D. simulans* Bam by 8 amino acids. We hypothesized that these sequence differences may be causing *D. simulans* Bam protein hyper-accumulation. To test this we generated the construct *sim-bam-melpest-yfp* that is identical to *sim-bam-yfp* except that we replaced the predicted *D. simulans* PEST domain with the predicted *D. melanogaster* Bam PEST sequence (Figure 3.2). We then generated *D. melanogaster bam+* flies which carried four copies of the *sim-bam-melpest-yfp* transgene. We find that these flies still show ectopic Bam-YFP expression in egg chambers (Figure 3.17C). These data demonstrate that other changes in sim-Bam-YFP are likely altering Bam-YFP protein stability in females.

3.4 Discussion

As a class, reproductive proteins evolve at higher rates than most other protein classes. While population genetic analyses have provided important data about the prevalence, temporal nature, and history (across a phylogeny) of adaptive evolution, a more recent focus of the field has been to determine the functional consequences of protein adaptation (MacCallum and Hill, 2006; Jensen et al., 2007; Storz and Wheat, 2010; Storz and Zera, 2011). We have focused our study on the adaptive evolution of *bam* not only to identify the functional consequence of *bam*'s sequence divergence, but also to utilize this information to make hypotheses about what is driving *bam* to adaptively evolve which we ultimately aim to test.

Using an interspecies rescue assay we identified functional divergence of *bam*. We found that although *sim-bam-yfp* complements *D. melanogaster bam* male

sterility, it fails to fully complement female sterility, demonstrating that *bam* has functional diverged in its role in female fertility.

3.4.1 Transgenesis Rescue Assay

We created an interspecies rescue assay to address the functional divergence of Bam's amino-acid substitutions. We generated *D. melanogaster* and *D. simulans bam* constructs driven by the *D. melanogaster bam* regulatory region. The regulatory region of *bam* was defined from constructs that rescued both male and female sterility (McKearin and Spradling, 1990). Further studies more finely mapped the 5' region of *bam* needed to silence its expression in GSCs and also tested the importance of the *bam* ORF and 3' region in its regulation (Chen and McKearin, 2003b; Song et al., 2004). These experiments found that transgenes with approximately 200 bp upstream of the *bam* 5'UTR have full rescuing activity, and that both the *bam* ORF and 3'UTR are dispensable for proper transcriptional regulation. Using quantitative RT-PCR, we find that using an upstream regulatory region of 1.7 kb is insufficient to drive wildtype *bam* expression in *mel-bam-yfp;bam*$^{-}$ ovaries. This result was surprising, but further expression analysis showed it is not due to genetic background effects of either the transgenic line or the heterozygous control used for comparison (Figure 3.4). We performed qRT-PCR on the original *bam* rescue transgene, *bam-α;bam*$^{-}$ (McKearin and Spradling, 1990) and found that it too is underexpressed. These data suggest that the sequences necessary to drive wildtype expression of *bam* in females are not included in either transgene. The initial mapping studies of the *bam* regulatory region were performed in a qualitative manner, looking for proper RNA or protein expression

patterns by in situ or IHC, and rescue was assayed qualitatively by looking for fertile males and females. Our data suggest that although *bam* RNA levels in *mel-bam-yfp;bam*$^{-}$ are underexpressed, its protein expression level is more similar to *sim-bam-yfp;bam*$^{-}$. Since *bam* RNA levels in *sim-bam-yfp;bam*$^{-}$ ovaries is the same as the *D. melanogaster bam* heterozygote, we presume, although have not been able to test, that Bam protein levels in *sim-bam-yfp;bam*$^{-}$ flies are similar to that of the *D. melanogaster bam* heterozygote as well. Therefore, it is likely that the qualitative expression assays used to map the regulatory region would not detect these smaller differences, as *bam* would be expressed in the correct pattern. Our data also show that this reduction in *bam* expression does not affect the ability of *mel-bam-yfp* to rescue female sterility. Furthermore, we only find lack of rescue in males under sperm exhaustion conditions, which qualitative assays are unlikely to detect. I have thus far been unable to measure Bam-YFP protein levels in testes to compare to RNA levels, as they are expressed too low to detect by Western Blot. One interesting possibility is that the incongruence between RNA and protein levels may indicate a feedback mechanism that acts on *bam* at the level of translation. If this were true, it would suggest that there is a feedback mechanism in females but not in males, which is why we see male fertility defects.

The cause of the different RNA levels between *mel-bam-yfp;bam*$^{-}$ and *sim-bam-yfp;bam*$^{-}$ is unclear. Because they have the same non-coding DNA from *D. melanogaster*, the difference must lie within the CDS (or small introns). It may be that each transgene is transcribed the same and that *sim-bam-yfp* RNA levels are more stable. Alternatively, *sim-bam-yfp* and *mel-bam-yfp* may be transcribed at different

rates. For example, differences in the secondary structures of the transgenes may hinder the ability of polymerase to transcribe, or coding sequence elements may stimulate promoter or enhancer activity. Both examples have been have been implicated in transcriptional regulation (Neznanov et al., 1997; Morel and Massoulie, 2000). Although we cannot distinguish between these explanations, the end result is that *bam* levels in *sim-bam-yfp;bam*$^{-}$ are expressed similarly to the *D. melanogaster bam* heterozygote. Thus, our study further compared *sim-bam-yfp;bam*$^{-}$ to the *D. melanogaster bam* heterozygote to be able to attribute the functional differences we detected to coding sequence divergence rather than expression differences.

3.4.2 Bam's role in female fertility has diverged but is not due to an inability of sim-Bam to interact with mel-Bgcn

We found that *sim-bam-yfp* does not rescue female sterility to the level of either *mel-bam-yfp* or the *D. melanogaster bam* heterozygous control, while *sim-bam-yfp* fully rescues male sterility. In a more stringent assay in which the males were sperm exhausted, we found that *sim-bam-yfp* could still rescue male sterility, suggesting that *D. simulans bam* has not functionally diverged with respect to male fertility. It is possible that there could be subtle fertility defects in *sim-bam-yfp;bam*$^{-}$; however, relative to the large differences we see in females, these differences would be extremely small. Thus, our transgenesis assay system has shown evidence for functional divergence between the *D. melanogaster* and *D. simulans bam* orthologs and that this divergence predominantly affects *bam*'s role in female fertility.

sim-bam-yfp;bam^{-} females have multiple ovarian defects. Since Bam and its interacting partner Bgcn are both adaptively evolving, we hypothesized that these ovarian defects might be due to an inability of sim-Bam to interact with mel-Bgcn. We provide three lines of evidence against this: 1) *bgcn* is required for *bam*'s role in GSC differentiation. If this interaction were eliminated or reduced, we would expect to see tumorous ovaries but never do in *sim-bam-yfp;bam*$^{-}$ flies, 2) sim-Bam::HA and mel-Bgcn::MYC reciprocally co-immunoprecipitate with one another in S2 cells, and 3) removing one copy of *bgcn* does not exacerbate *sim-bam-yfp;bam*$^{-}$ ovarian defects nor does it cause tumorous ovaries. This combination of biochemical and genetic data strongly suggest that the interaction between sim-Bam and mel-Bgcn is maintained and that the *sim-bam-yfp;bam*$^{-}$ defects are due to other incompatibilities.

Why does *bam*'s functional divergence only affect female fertility and not male fertility? This may seem implausible given *bam*'s similar roles in males and females; however previous studies suggest that not all *bam* functions are shared between males and females and that the sexes have different sensitivities to *bam* levels. Forced expression of a *bam* transgene in GSCs resulted in their differentiation only in females and not males (Ohlstein and McKearin, 1997). Only after males were exposed to a longer duration and occurrence of heat shock were GSCs lost in males (Kawase et al., 2004). Additionally, only in females does *bam* function primarily in GSC differentiation. Instead, *bam*'s primary role in males is in regulating cyst divisions and entry into meiosis. Elegant studies have shown that increased or decreased levels of *bam* result in cysts with either less or more cells per cyst, respectively, which presumably can still give rise to mature sperm (S. DiNardo, pers.

comm.). Therefore, we believe that changes in how *bam* functions in female fertility will have minimal effects on its role in males, as they only require *bam* for regulation of cyst divisions and this function is more permissive to changes in *bam* expression.

3.4.3 The molecular nature of *sim-bam-yfp;bam*$^{-}$ defects

We have shown that *sim-bam-yfp* is more abundant than *mel-bam-yfp* and causes a more severe phenotype with increased copy number. To determine if mel-Bam-YFP is underexpressed at the protein level or if sim-Bam-YFP is in fact overexpressed relative to the *D. melanogaster bam* heterozygote, we are currently using antibodies generated to Bam. We hope these additional studies can help narrow down the molecular nature of *sim-bam-yfp;bam*$^{-}$ defects.

While we currently do not know the molecular nature of *sim-bam-yfp* defects, our dominance studies provide valuable insight to this question. Increasing *sim-bam-yfp* dose results in increased GSC loss which likely results from cytoplasmic sim-Bam-YFP within the GSC. While *bam* is transcriptionally repressed in the GSC in wildtype *D. melanogaster*, there is a small amount of Bam protein present in the GSC which must be kept inactive (Jin et al., 2008). Two different hypotheses of how Bam is kept inactive in GSC have been proposed. The first hypothesis is based on data in which a subset of antibodies generated to Bam show that it is localized at the fusome (McKearin and Ohlstein, 1995). Additionally, *bam* is required for proper fusome formation. In *bam* mutants, Ter94, a known fusome component, does not properly localize (Leon and McKearin, 1999), and the fusome shows a reduced amount of vesicular material (McKearin and Ohlstein, 1995). Under this hypothesis, all Bam

protein present in the GSC is localized at the fusome, rendering it inactive. A second hypothesis suggests that there is indeed a small amount of cytoplasmic Bam present in the GSC, but that other proteins antagonize its activity (Jin et al., 2008). Only after Bam accumulates to high levels can it titrate away antagonizing proteins and bind to other partners to promote differentiation (Xie, T. 2010 Drosophila Research Conference). Based on these data, we hypothesize that *sim-bam-yfp;bam*$^{-}$ GSC loss could result from multiple problems that are not necessarily mutually exclusive: 1) sim-Bam-YFP is overexpressed in the GSC, 2) Proteins that are necessary to antagonize Bam cannot properly interact with sim-Bam-YFP, 3) Bam participates in a feedback loop to restrict its GSC expression that *sim-bam-yfp* is unable to fully participate or, 4) sim-Bam-YFP cannot properly localize to the fusome and is cytoplasmically active. We can eliminate most of these possibilities based on the results of the dominance study which shows that the presence of *D. melanogaster bam* rescues *sim-bam-yfp;bam*$^{-}$ GSC loss. In hypotheses 1-2, it is unlikely that the presence of wildtype *bam* would cause a reduction in the expression of sim-Bam-YFP or allow for proteins to gain the ability to antagonize sim-Bam-YFP. In hypothesis 3, *sim-bam-yfp* would be unable to feedback to itself, but the presence of wildtype *bam* could provide the necessary feedback to regulate *sim-bam-yfp*. We have measured protein levels by Western blot and find no evidence of a decrease in sim-Bam-YFP in the presence of wildtype *D. melanogaster* Bam (Figure 3.18). However, it is important to note that westerns were performed on whole ovaries and not just GSCs. In hypothesis 4, sim-Bam-YFP is unable to properly form the fusome, and therefore, is unable to localize to the fusome. The presence of *D. melanogaster bam* allows for proper

formation of the fusome and subsequent localization to it by sim-Bam-YFP. We favor this hypothesis since *sim-bam-yfp* flies also show mitotic synchrony defects, a hallmark of improper fusome function. Moreover, proper endocytic recycling of the fusome is required for GSC maintenance, as *rab11* mutants show GSC loss and have defects similar to *bam* mutants (Lighthouse et al., 2008). We are interested in looking at Bam-YFP localization at fusomes; however, the Bam-F antibody that recognizes fusome-associated Bam is no longer available (D. McKearin, pers. comm.). It would be interesting to look at fusome morphology by electron microscopy to determine if any defects are present. However, we have not had success with this method.

3.4.4 Selective pressures acting on *bam* in females

Our transgenic rescue assay has shown that *bam* has diverged specifically in regard to its role in female fertility, suggesting that the selective pressure driving the adaptive evolution of *bam* is likely acting in the female germline. Below we explore female germline pressures and their potential to drive the adaptation of *bam*.

One hypothesis is that *bam* is evolving in response to a novel environment or change in environment, resulting in a change in reproductive pressures. *D.*

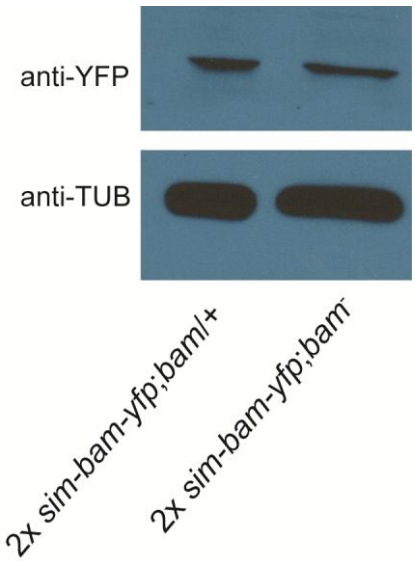


Figure 3.18: The addition of wildtype *D. melanogaster bam* does not reduce sim-Bam-YFP protein levels. Western blot probed with antibodies against anti-YFP or anti-α-Tubulin. This experiment was performed one time. Afterward it was found that a single *bam* mutant stock was contaminated. I do not know if contamination may have affected these results.

melanogaster and *D. simulans* are believed to have evolved in allopatry (Dean and Ballard, 2004). This hypothesis would require each species to experience ecological pressures over a similar time period and respond to these challenges by changing the same gene. No thorough studies have been done comparing the germline progression between these two species, but those that have find little differences in germline development (Hollocher et al., 2000), and we find no differences in the ability of Bam or a few other germline markers to properly localize in females of *D. simulans, D. sechellia,* and *D. yakuba* using IHC (Figure 3.19). Therefore, we believe this is an unlikely scenario although a more systematic comparison of the protein expression/localization and cell types would be needed to fully eliminate this scenario.

Alternatively, *bam* could be evolving in response to conflict in the female germline. The adaptive evolution of many reproductive genes is attributed to sexual conflict and/or sexual selection. This is also an unlikely explanation for *bam* as it is expressed earlier in reproduction, limiting its opportunity to interact with proteins of the opposite sex. Many genes involved in the repression of transposable elements (TEs) have also been shown to evolve under adaptive evolution (Obbard et al., 2009; Kolaczkowski et al., 2011) and many of these proteins act primarily in the germline to silence TEs via the piRNA pathway (reviewed in Aravin et al., 2007; Senti and Brennecke, 2010). While piRNA genes are present in both males and females to repress TEs, maternal contribution is particularly important to feedforward this protection to offspring (Brennecke et al., 2008). It is unlikely that *bam* plays a role in

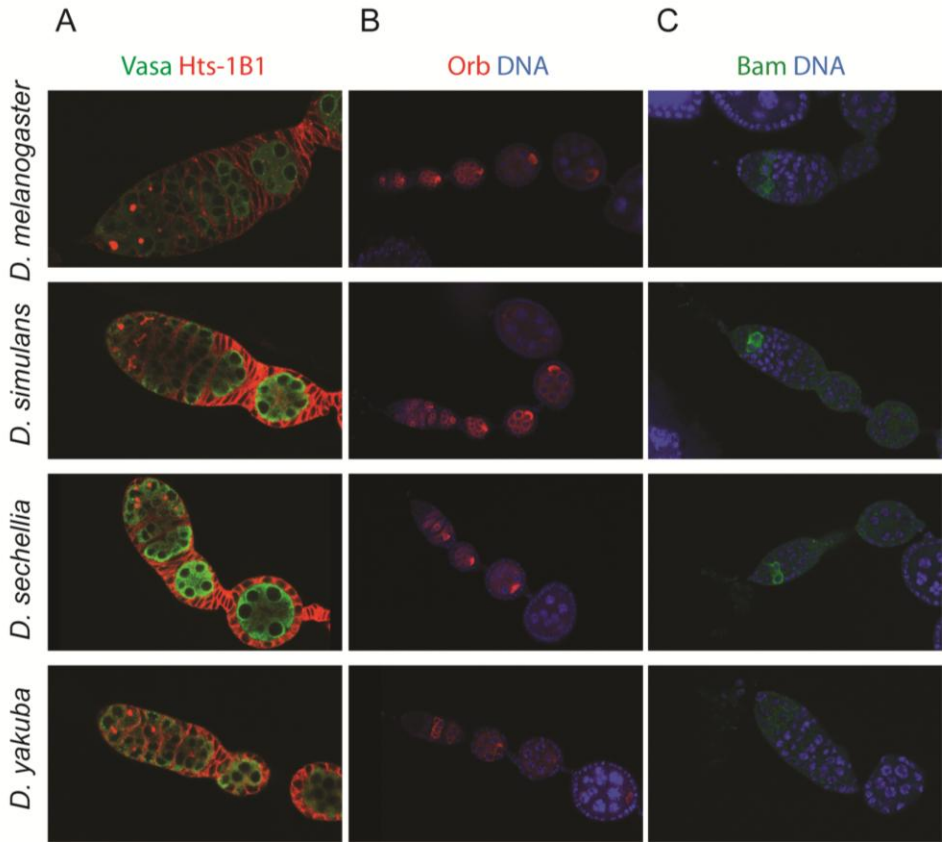


Figure 3.19: Germline markers localize similarly across different Drosophila species. Ovaries from *D. melanogaster*, *D. simulans*, *D. sechellia*, and *D. yakuba* are stained with antibodies to Vasa and Hts-1B1 (column A), Orb (Column B), or Bam (Column C). Anti-Bam does not cross-react with *D. yakuba* Bam.

TE repression, however, as a recent study has compared the small RNA production and transposon regulation between *bam* mutants and wildtype controls and find no discernable differences (Rangan et al., 2011).

Alternatively, GSC competition could be driving the adaptive evolution of *bam*. GSCs attach to the niche through the expression of E-cadherin at the GSC-niche interface (Xie and Spradling, 2000; Song et al., 2002). *bam* and *bgcn* mutant GSCs accumulate more E-cadherin at this interface and eventually out-compete wildtype stem cells for niche occupancy. *bam* overexpression is sufficient to downregulate E-cadherin and it has been proposed that this regulation is a quality-control mechanism to force defective GSCs out of the niche (Jin et al., 2008). While mutations in *bam* could alter a GSC's competitiveness relative to a wildtype stem cell, the competitive advantage would last for only a single generation. When the progeny generated from mutated stem cells establish their germline, all GSCs would again be clonal. At this point, genetic drift would be the only selective force acting on GSCs. Although the selective advantage over a single generation would, by definition, increase the frequency and, therefore, the probability of fixation of the particular mutation, overall the frequency of the mutation would still be incredibly low and subject to strong drift. Therefore, it seems unlikely that stem cell competition would drive recurrent substitutions in *bam.*

Wolbachia pipientis is an intracellular bacterial endosymbiont that is maternally inherited. It elicits a number of reproductive manipulations to maximize its transmission including male killing, feminization, parthenogenesis, and cytoplasmic incompatibility (reviewed in Werren et al., 2008). *Wolbachia* can be nearly or

absolutely essential for host fertility (Dedeine et al., 2001; Fast et al., 2011). *Wolbachia* can also partially rescue the fertility of *D. melanogaster Sex lethal* (*Sxl*) mutants (Starr and Cline, 2002). This result is intriguing given that *Sex lethal* is required for *bam* to function in GSC differentiation (Chau et al., 2009). Perhaps *bam* too interacts with *Wolbachia* and this interaction is driving *bam* to adaptively evolve. The nature of the interaction between *Wolbachia* and *Sxl* is unknown, but it is possible that *Wolbachia* may be manipulating host genes to promote GSC differentiation rather than symmetric replacement. In this scenario, mutations in *bam* may be under selection as they will re-stabilize the GSC population. Alternatively, *Wolbachia* may utilize host factors to determine where to localize or how often to replicate. In this scenario, mutations in *bam* may be selected if they do not allow for efficient *Wolbachia* transmission.

Studies have shown that many male reproductive proteins are adaptively evolving, suggesting that strong selective pressures are present in males (Aguade et al., 1992; Begun et al., 2000; Andres et al., 2006; Wong et al., 2008; Li et al., 2009). My study provides evidence that not only are there likely selective pressures acting in the female germline, but that they act at very early stages of germline development. Using the well-studied system of the Drosophila germline, we tested for functional divergence of the *bag of marbles* gene, moving from the adaptive genotype to identifying the adaptive phenotype. Our data show that *bam* has functionally diverged specifically in regard to its role in the female germline, which has been important in generating informed hypotheses about the selective pressure driving the adaptation of *bam* which we can now test.

CHAPTER 4

WOLBACHIA GENETICALLY INTERACTS WITH THE DROSOPHILA *BAG OF MARBLES* GENE

4.1 Introduction

The field of developmental biology is focused on questions of how genes function in the development of an organism, such as the tissues in which they are expressed, the genes with which they interact, and the pathways in which they participate. Many of the genes involved in these critical processes are evolutionarily conserved both in sequence and in their presence across a broad range of taxa. However, some of these critical regulators show high rates of amino acid substitutions driven by positive selection. It may seem counterintuitive for proteins that play these critical roles to evolve so fast, but there are many selective pressures acting on an organism. Host reproduction, in particular, is a process in which a high proportion of genes have experienced adaptive evolution. A number of selective pressures have been suggested to drive the adaptive evolution of reproductive proteins including sexual conflict, sexual selection, pathogen defense, and avoidance of self-fertilization (Swanson and Vacquier, 2002; Clark et al., 2006; Panhuis et al., 2006).

The Drosophila ovary provides an important system in which to study host reproduction and the selective pressures that act on it. The ovary is comprised of units called ovarioles that are an assembly line for the developing oocyte. At the most anterior end of the ovariole is the germarium in which the germline stem cells (GSCs)

are present. The *bag of marbles* (*bam*) gene plays an important role in this system. *bam* is a differentiation promoting factor and is therefore repressed in the GSCs (McKearin and Ohlstein, 1995; Chen and McKearin, 2003a; Chen and McKearin, 2003b). However, when a GSC daughter moves even one cell-width away it will express *bam* and start to differentiate (reviewed in Wong et al., 2005; Kirilly and Xie, 2007). This daughter cell will undergo four rounds of mitotic division in syncytium, giving rise to a 16-cell cyst, one cell of which will become the future oocyte. *bam*'s role in GSC differentiation requires the function of at least two other genes, *benign gonial cell neoplasm (Bgcn*) (Lavoie et al., 1999; Ohlstein et al., 2000) and *Sex Lethal* (*Sxl*) (Chau et al., 2009).

We and others have previously shown that the *bag of marbles* gene has diverged under rapid, adaptive evolution in both *D. melanogaster* and *D. simulans* (Civetta et al., 2006; Bauer DuMont et al., 2007). In Chapter 3, we generated a transgenic assay to identify the functional consequences of *bam*'s sequence divergence. We assayed the ability of a *D. simulans bam* ortholog to rescue the male and female sterility of a *D. melanogaster bam* mutant. While *D. simulans bam* was able to rescue male sterility, it was unable to fully rescue the female sterility of *D. melanogaster bam* mutants, suggesting that the selective pressure acting on *bam* is in the female germline. I discussed some of the selective pressures that are present in the female germline in Chapter 3. In this Chapter, I test one of these hypotheses; that the bacterial endosymbiont, *Wolbachia pipientis*, has driven *bam* to adaptively evolve. *Wolbachia* is maternally inherited and has been shown to manipulate host reproduction in a variety of organisms (Veneti et al., 2005; Hornett et al., 2006; Serbus

et al., 2008; Werren et al., 2008). Most strikingly, *Wolbachia* infection can rescue the oogenesis defects of *Sxl* mutants (Starr and Cline, 2002; Sun and Cline, 2009), and a recent study has shown that *bam* requires *Sxl* to function in GSC differentiation (Chau et al., 2009).

Wolbachia localization and titer are controlled by both the bacteria and the host. In *D. melanogaster*, *Wolbachia* is present throughout the germline of the females but preferentially accumulates at the somatic stem cell niche (Frydman et al., 2006; Fast et al., 2011). The somatic stem cell niche is a microenvironment required to maintain somatic stem cells that, when differentiated, will give rise to the follicle cell layer that surrounds egg chambers. In naturally occurring infections, the localization of *Wolbachia* in ovaries is dramatically different among closely related species. In *D. mauritiana*, *Wolbachia* preferentially to the germline stem cell niche (Fast et al., 2011). In *D. simulans*, *Wolbachia* localizes to both the somatic and germline stem cell niches (Frydman, H., 2012 Drosophila Research Conference Poster). Through a combination of transinfection assays and introgression studies, it has been shown that *Wolbachia*'s tissue tropism in the ovary is primarily encoded by the bacteria (Frydman, H., 2012, Drosophila Research Conference Poster, http://www.drosophila-conf.org/2012/book/Abstracts_Web.pdf).

Transinfection studies where one strain of *Wolbachia* is moved into a different species have been done between the *Wolbachia* strains that infect *D. melanogaster* and *D. simulans* (Boyle et al., 1993; Poinsot et al., 1998; McGraw et al., 2001; McGraw et al., 2002). A common reproductive manipulation by *Wolbachia* in these species is cytoplasmic incompatibility (CI). In CI, when *Wolbachia*-infected fathers mate with

uninfected mothers, *Wolbachia* induces the death of their offspring. In the transinfection experiments, one *Wolbachia* strain used was native to *D. melanogaster*, has lower titer levels, and did not induce high levels of CI in its native host. The second strain used was a *Wolbachia* that was native to *D. simulans*, has high titer levels, and induces high levels of CI. When CI-inducing *Wolbachia* from *D. simulans* were transferred to *D. melanogaster*, their ability to induce CI decreased dramatically (Boyle et al., 1993). Conversely, when strains that do not induce high CI in *D. melanogaster* were transinfected into *D. simulans*, they induced high levels of CI. This increase in CI is correlated with *Wolbachia* acquiring the ability to colonize sperm bundles in the new host and replicating to much higher levels (Poinsot et al., 1998; McGraw et al., 2001). These results show that strains from both species have the ability to induce CI but that host control over *Wolbachia*'s tissue tropism and replication in a species-specific manner likely reduces the induction of CI.

In this study we investigate the hypothesis that *Wolbachia* is the selective pressure driving the adaptive evolution of *bam* by testing for an interaction between *bam* and *Wolbachia*. Using a combination of *D. melanogaster bam* hypomorphs and orthologous *bam* transgenes, we find that the presence of *Wolbachia* can partially rescue their female sterility defects. The primary focus of this study was to determine the molecular nature of this interaction and to test models of how it may have driven *bam* to adaptively evolve.

4.2 Materials and Methods

4.2.1 Drosophila stocks

All stocks were cultured at room temperature on standard yeast-glucose media. The $bam^{\Delta 86}$and bam^{BW} (also referred to as $bam^{z3\text{-}2884}$ (Ohlstein et al., 2000) or bam^{Z} (Shen et al., 2009)) stocks are described in FlyBase. The $bam^{\Delta 59}$ allele is described in Chapter 3. These stocks were kindly provided by Dr. Dennis McKearin. The generation of the transgenic lines, *sim-bam-yfp* and *mel-bam-yfp*, is described in Chapter 3. The *w*Mel-infected strain of *D. melanogaster*, *w*; *Sp*/*CyO; Sb*/*TM6B* +*w*Mel, was kindly provided by Dr. Bill Sullivan.

4.2.2 *Wolbachia* infection and tetracycline treatment

We generated a $bam^{\Delta 59}$*/TM3* stock that was infected with the *w*Mel strain of *Wolbachia* from *D. melanogaster* by crossing *w*Mel-infected females into the *bam* mutant stock, generating $bam^{\Delta 59}$*/TM3* +*w*Mel. The $bam^{\Delta 59}$*/TM3* +*w*Mel stock was then cured of *Wolbachia* by feeding the flies on media supplemented with 0.03% tetracycline for three generations, generating $bam^{\Delta 59}$*/TM3* Tet. $bam^{\Delta 59}$*/TM3* +*w*Mel females were then backcrossed to the $bam^{\Delta 59}$*/TM3* Tet strain for at least six generations to generate genetically similar backgrounds including the mitochondria.

4.2.3 Bam genetic interaction assays

A combination of hypomorphic *bam* alleles has been used to identify genetic interactions with *bam* (Ohlstein et al., 2000; Shen et al., 2009). The *bam* hypomorphs

give rise to a mix of tumorous and wildtype egg chambers and the females are weakly fertile (Ohlstein et al., 2000). By measuring the number of nurse-cell positive egg chambers per ovariole, one can identify suppressors or enhancers of *bam* activity. To assay the effect of *Wolbachia* infection of *bam*, we crossed females of either $bam^{\Delta 59}$*/TM3 +w*Mel or $bam^{\Delta 59}$*/TM3* Tet to males of the genotype bam^{BW}*/TM3*. Non-*TM3* female progeny with and without *Wolbachia* were used in fertility assays or for ovary examination. It is important to note that *Wolbachia*'s ability to induce cytoplasmic incompatibility is dependent on male age and male developmental time, termed the "younger brother" effect (Yamada et al., 2007). It is unknown whether these effects are also important on other *Wolbachia*-host interactions or whether any "younger sister" effect occurs in females. My experiments controlled for post-eclosion age but not pre-eclosion age, and are, therefore, not controlled for any potential "younger sister" effect.

4.2.4 Immunostaining

Immunostaining was performed as in Aruna *et al*. (2009). The following primary antibodies were used: anti-Vasa (DSHB, 1:25), anti-Hsp60 (Sigma H3524, 1:100). Secondary antibodies including goat anti-rat, anti-rabbit, or anti-mouse were conjugated with Alexa 488 or 546 (Molecular Probes, 1:200-1:500). Samples were mounted in Vectashield containing DAPI (Vector Laboratories) and analyzed using the Leica SP2 confocal microscope at the Cornell University Core Life Sciences Microscopy and Imaging Facility. Images were resized in Photoshop (Adobe, version 11.0).

4.2.5 Fertility assays

All crosses were performed at room temperature (22-23°C). Prior to crossing all virgin flies were aged for 2-3 days post-eclosion on media supplemented with yeast. In female fertility experiments, single transgenic females were crossed to two wildtype Canton S (CS) males. The trio of flies were transferred to a new vial every five days for a total of 15 days and then discarded. Progeny from each vial were counted for 8 days after eclosion. In male fertility experiments, single males were mated to two wildtype CS females as described above.

4.2.6 Quantitative RT-PCR

Females were aged for 2-3 days post-eclosion on media supplemented with yeast, and total RNA was isolated from 10 ovaries using Trizol reagent (Invitrogen) following the manufacturer's protocol. Samples were processed as described in Chapter 3. For *bam*, primer pair 1160/1170 (Table 4.1) amplified *bam* from both species with high efficiencies. For *rpl32*, primer pair 844/845 (Table 4.1) from Maheshwari and Barbash (2012) was used. The standard curve method was used to estimate *bam* and *rpl32* levels. Three technical replicates were performed from at least three biological replicates for each sample.

4.2.7 Quantitative PCR

To assay levels of *Wolbachia*, qPCR was performed on genomic DNA as in (McGraw et al., 2001; McGraw et al., 2002). Females were aged for 2-3 days post-eclosion on media supplemented with yeast. DNA was isolated from ovaries using

phenol-chloroform extraction. 3-10 ovaries were homogenized in 50 ul of homogenization buffer (10 mM Tris-HCl pH7.5, 60 m M NaCl, 10 mM EDTA, 0.15 mM Spermidine, 5% Sucrose). 50 ul of lysis buffer (200 mM Tris-HCl pH9.0, 30 mM EDTA, 2% SDS, 5% Sucrose) and 2 ul of proteinase K (10 mg/ml) were added, followed by incubation at 37°C for 2 hours. The solution was extracted twice with 100 ul Phenol:Chloroform:Isoamyl alcohol (25:24:1) and then extracted with an equal volume of chloroform. NaCl was added to a final concentration of 0.2 M along with twice the volume of 100% ethanol and the DNA precipitated for at least 1 hour at -20°C. DNA was precipitated by centrifugation at 14,000 rpm for 10 minutes at 4°C and washed with 70% ethanol. The DNA was rehydrated in TE.

The samples were further purified adding 0.1 volumes of 7.5 M NH4OAc and 2.5 volumes of 95% Ethanol. Samples were incubated on ice for 10 minutes and centrifuged at 14,000 rpm for 10 minutes at room temperature. The ethanol was removed and samples were washed with 70% ethanol. Samples were desiccated under a SpeedVac and hydrated in water.

For *Wolbachia*, primer pair wsp440F/wsp691R generated to the *wsp* gene was used (McGraw et al., 2001). For *rpl32,* primer pair 844/845 was used (Table 4.1). The standard curve method was used to estimate levels of each product. Three technical replicates were performed from at least three biological replicates for each sample.

Table 4.1: Primers used in this study

No.	Primer name	Sequence
844	RPL32-F	CCGCTTCAAGGGACAGTATC
845	RPL32-R	GACAATCTCCTTGCGCTTCT
1169	bam_rt_F3	GCCCATAACTATTGAGAAACTGC
1170	bam_rt_R3	GATCATGCAGGGATCTGAACAG
	Wsp440F	CTGGTGTTAGTTATGATGTAAC
	Wsp691R	AAAAATTAAACGCTACTCCA

4.3 Results

4.3.1 *Wolbachia* infection rescues the sterility of *bam* hypomorphic mutants

To determine if any interaction existed between *bam* and *Wolbachia*, we crossed a naturally occurring strain of *D. melanogaster Wolbachia*, *w*Mel, into a heteroallelic combination of *bam* hypomorphic alleles that has been used in genetic-interaction assays to look for enhancers or suppressors of *bam* (Ohlstein et al., 2000; Shen et al., 2009). $bam^{BW}/bam^{\Delta 59}$ flies lacking *Wolbachia* are weakly fertile, giving rise to a mix of tumorous and wildtype egg chambers (Ohlstein et al., 2000). The number of nurse-cell positive egg chambers per ovariole was compared between $bam^{BW}/bam^{\Delta 59}$ flies infected *w*Mel, denoted as "*bam* hypomorph +*w*Mel" and hypomorphic flies that had been cured of *Wolbachia* using tetracycline, denoted as "*bam* hypomorph Tet". We find that the ovarioles of *bam* hypomorph +*w*Mel flies contain significantly more nurse-cell-positive egg chambers than the *bam* hypomorph Tet flies (Table 4.2).

We then assayed the fertility of the *D. melanogaster bam* hypomorphic flies with and without *Wolbachia*. We find that the presence of *Wolbachia* increases the fertility of the *bam* hypomorph females to near wildtype levels as seen in other fertility experiments (Chapter 3; P. Satyaki, pers. comm.; E. Kelleher-Meisel, pers. comm.) (Figure 4.1). The fertility increase was only observed in *bam* hypomorphs and not in combinations of *bam* null alleles that result in complete sterility of females ($bam^{\Delta 86}/bam^{\Delta 59}$+*w*Mel N=20). Fertility assays were also performed in males. However, *bam* hypomorphic males were completely sterile, and the presence of

Table 4.2: *Wolbachia* genetically interacts with *D. melanogaster bam*

Genotype	# of nurse cell-positive egg chambers	# of ovarioles examined
D. melanogaster bam hypomorph Tet	372	126
D. melanogaster bam hypomorph +*w*Mel	350	190

F.E.T. P = 9.5e-4

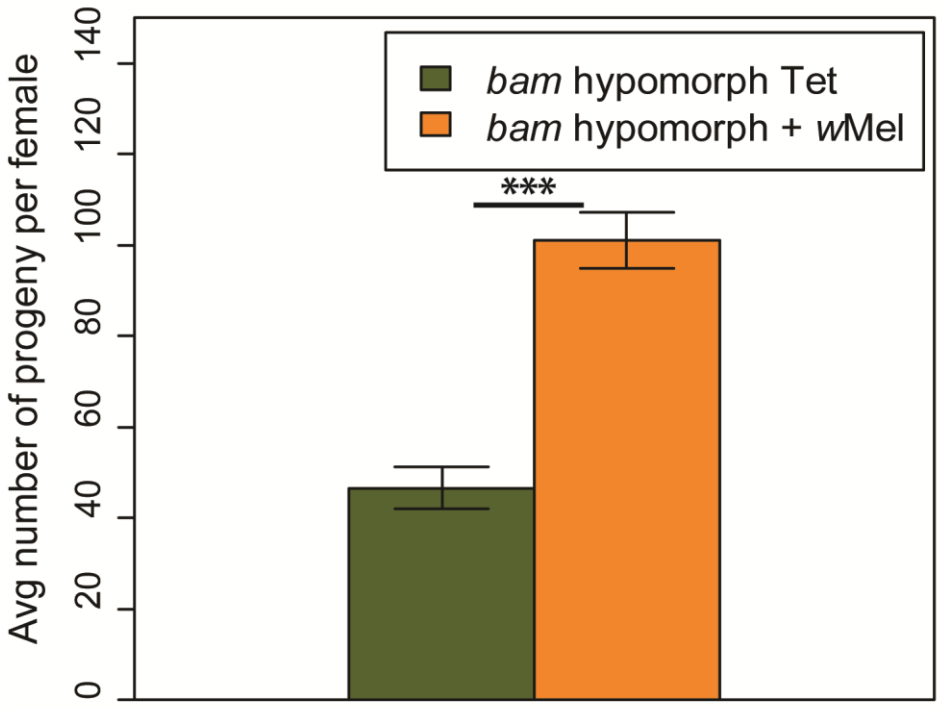


Figure 4.1: *Wolbachia* increases the fertility of *D. melanogaster bam* hypomorphs. One female and two tester males were allowed to mate and the trio was removed from the vial after 8 days. Fertility is shown as the average number of progeny per female +/- SEM for each vial. *Wolbachia*-infected *bam* hypomorphs are significantly more fertile than uninfected *bam* hypomorphs (T-test, ***P<0.001).

Wolbachia had no rescuing effect (*bam* hypomorph +*w*Mel N= 20, *bam* hypomorph Tet N=20).

4.3.2 *Wolbachia* increases the fertility of *D. melanogaster bam* mutants with *sim-bam-yfp*

In Chapter 3, we described a transgenic assay to identify functional divergence between *bam* orthologs. We generated YFP-tagged transgenes containing the *D. melanogaster* (*mel-bam-yfp*) or *D. simulans* (*sim-bam-yfp*) *bam* sequence driven by the *D. melanogaster bam* regulatory region. We then compared the ability of the *bam* orthologs to rescue the sterility of *D. melanogaster bam* mutant males and females and found that *sim-bam-yfp* rescues male sterility but failed to fully rescue female sterility. We refer to *mel-bam-yfp* or *sim-bam-yfp* in a *D. melanogaster bam* null background as *mel-bam-yfp;bam*$^{-}$ and *sim-bam-yfp;bam*$^{-}$, respectively. We compared the female fertility of each transgenic line with and without *Wolbachia* (*w*Mel) and found that the fertility of *mel-bam-yfp;bam*$^{-}$ was neither enhanced nor diminished in the presence of *Wolbachia*. However in comparisons of fertility in *sim-bam-yfp;bam*$^{-}$ females with and without *Wolbachia*, we found a significant increase in the fertility of young flies (Figure 4.3, Days 1-5) infected with *Wolbachia*. This result is reproducible among multiple insertion-site transgene pairs (Figure 4.4). These results provide additional, independent evidence that *bam* and *Wolbachia* genetically interact.

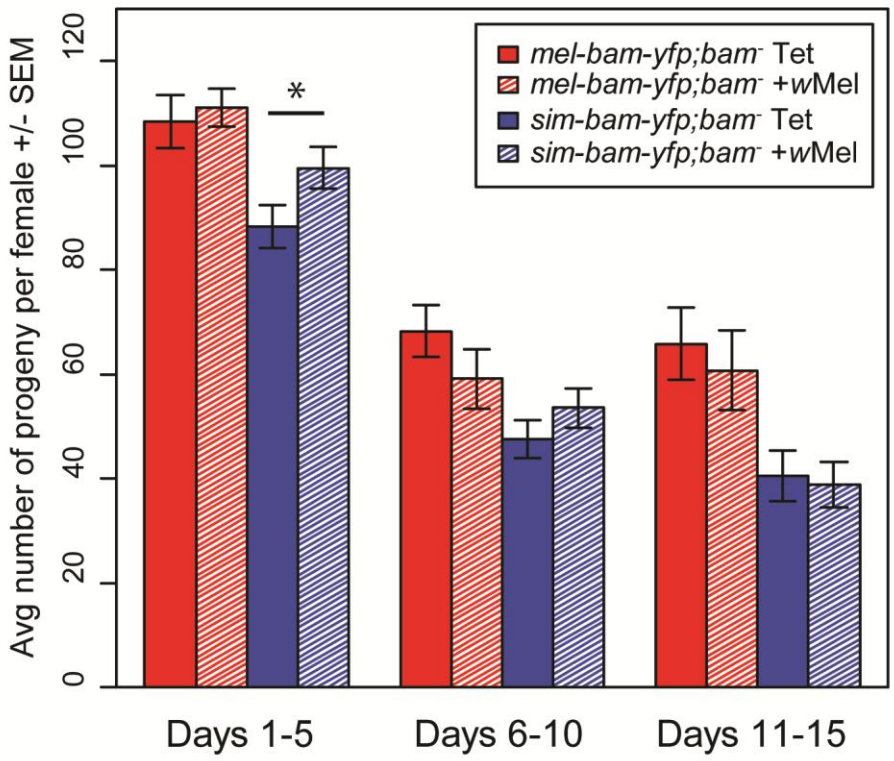


Figure 4.2: *Wolbachia* increases the fertility of young *sim-bam-yfp;bam*$^{-}$ females. One female and two tester males were allowed to mate and the trio was transferred to a new vial every five days. Fertility is shown as the average number of progeny per female +/- SEM for each vial. (T-test, *P<0.05). All comparisons between *mel-bam-yfp;bam*$^{-}$ +*w*Mel and *mel-bam-yfp;bam*$^{-}$ Tet are not significant. All other comparisons between *sim-bam-yfp*;*bam*$^{-}$ +wMel and *sim-bam-yfp;bam*$^{-}$ Tet (Days 6-10 and 11-15) are not significant. Transgenes are integrated in attP16a.

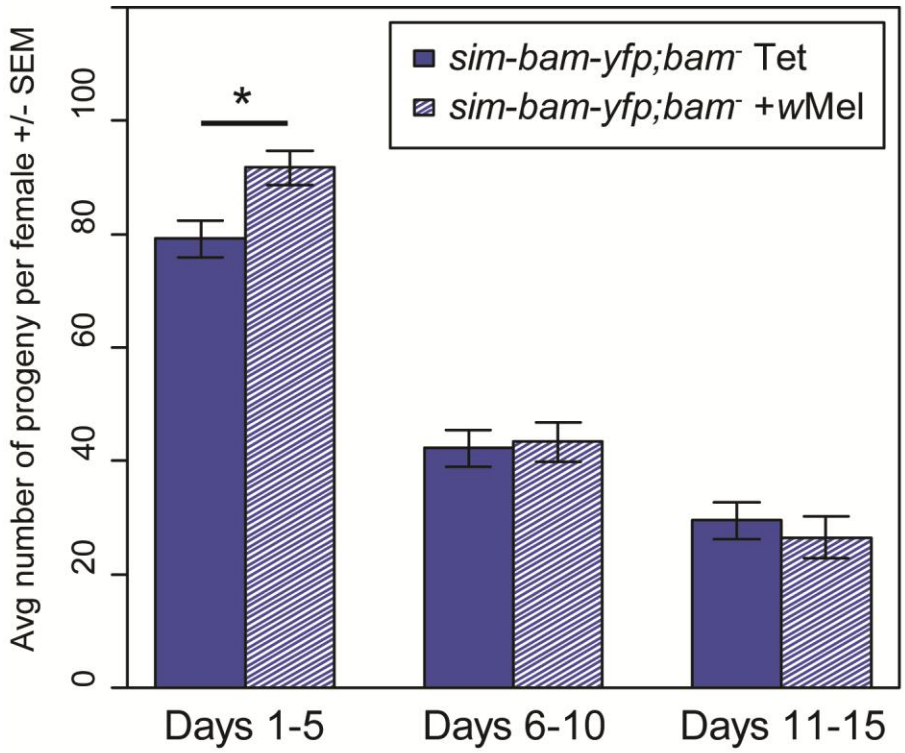


Figure 4.3: *Wolbachia* increases the fertility of *sim-bam-yfp;bam*$^{-}$ females (attP40). One female and two tester males were allowed to mate and the trio was transferred to a new vial every five days. Fertility is shown as the average number of progeny per female +/- SEM for each vial. Transgenes are integrated in attP40. (T-test, *P<0.05).

4.3.3 *Wolbachia*-mediated *bam* rescue is not due to an increase in *bam* RNA levels

Our genetic and fertility assay results suggest that *Wolbachia* infection may increase *bam* activity. If *Wolbachia* were acting directly on *bam* to increase its activity, one possible mechanism is via an increase in its RNA or protein level. We assayed the levels of *bam* mRNA in *D. melanogaster bam* hypormophic flies with and without *Wolbachia*. We find that *bam* hypomorph-Tet flies have RNA levels approximately half of what is found in flies wildtype for *bam*. This is not surprising as one of the *bam* alleles deletes most of the locus (Chapter 3). When we compare *bam* expression between *bam* hypomorph Tet and *bam* hypomorph +*w*Mel flies, we find no significant difference in the levels of *bam* RNA (Figure 4.4).

Many genes involved in GSC regulation show evidence of translational repression (Kadyrova et al., 2007; Sheng et al., 2009; Harris et al., 2011; Li et al., 2012a). Moreover, *Wolbachia* has been shown to alter the expression of genes at the post-transcriptional level (Hussain et al., 2011). Therefore, we also will need to compare protein levels between the *bam* hypomorphs with and without *Wolbachia*. I am currently optimizing conditions for an anti-Bam antibody to test this.

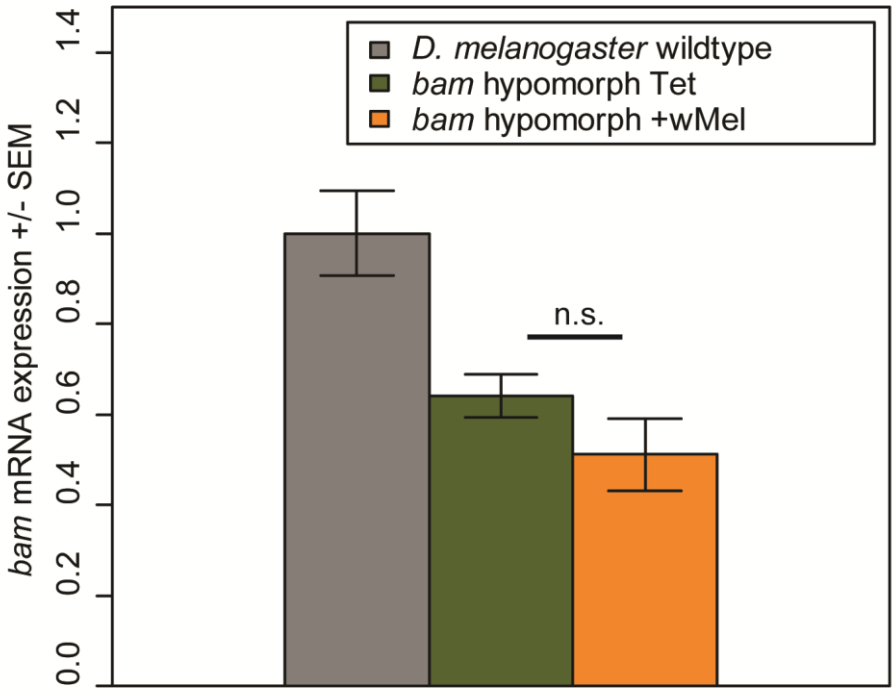


Figure 4.4: *Wolbachia* infection does not significantly alter *bam* RNA levels. qRT-PCR of the *D. melanogaster bam* hypomorph with and without *Wolbachia*. *bam* RNA levels in ovaries from the wildtype *D. melanogaster* strain, *yw* (grey), *bam* hypomorph Tet (green), and the *bam* hypomorph +*w*Mel (orange). There is no statistical difference in *bam* expression of the *bam* hypomorph with and without *w*Mel (T-test, P>0.05).

4.3.4 *D. melanogaster Wolbachia* titer is reduced in the presence of *sim-bam-yfp*

If *Wolbachia* has co-evolved with *bam*, one prediction is that *Wolbachia* would be more successful in the presence of its species-specific form of *bam*. To test this, we used qPCR to measure *Wolbachia* titer in ovaries to determine if *Wolbachia*'s ability to replicate changed in the presence of *sim-bam-yfp* relative to *mel-bam-yfp*. We find that *Wolbachia* levels are reduced in *sim-bam-yfp;bam*$^{-}$ ovaries compared to the levels in *mel-bam-yfp;bam*$^{-}$ ovaries (Figure 4.5). We have previously shown that *mel-bam-yfp* functions like wildtype *D. melanogaster bam* in ovaries. However, *bam* RNA expression levels in *mel-bam-yfp;bam*$^{-}$ ovaries is reduced relative to *bam* expression levels a *D. melanogaster bam* heterozygote, while *bam* expression in *sim-bam-yfp;bam*$^{-}$ is similar to *bam* expression levels in a *D. melanogaster bam* heterozygote (Chapter 3). To eliminate the possibility that this reduction in *Wolbachia* titer was due to difference in *bam* expression, we also want to compare *Wolbachia* titer in *sim-bam-yfp;bam*$^{-}$ ovaries to *D. melanogaster bam* heterozygous ovaries in the future.

4.3.5 *Wolbachia* localizes properly in the presence of *sim-bam-yfp*

While *Wolbachia* is present in low levels throughout the germarium, it preferentially accumulates in the somatic stem cell niche (SSCN) in *D. melanogaster* (Frydman et al., 2006; Fast et al., 2011). The SSCN replicates at rates much lower than the surrounding tissues, and it has been hypothesized that *Wolbachia* may preferentially localize here to allow for accumulation to high levels within the SSCN.

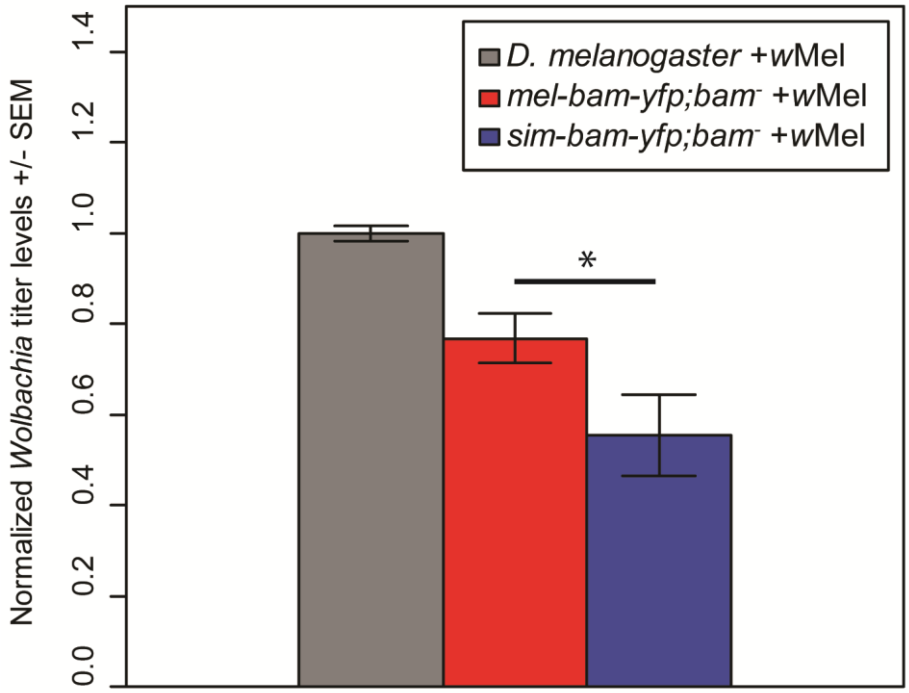


Figure 4.5: *D. melanogaster Wolbachia* cannot accumulate as well in the presence of *sim-bam-yfp*. q-PCR was performed using *Wolbachia*-specific primers to assay levels of *Wolbachia* present in the ovaries of *D. melanogaster* +*w*Mel (grey), *mel-bam-yfp;bam*$^{-}$ (red), *and sim-bam-yfp;bam*$^{-}$ (blue) flies. The *D. melanogaster* +*w*Mel line (*w; Sp*/*CyO*; *Sb*/*TM6B* +*w*Mel) was the line used to infect *bam* mutant stocks with *Wolbachia* and is wildtype for *bam*. (T-test, *P<0.05).

As germline cysts pass the SSCN, high *Wolbachia* titer would allow it to efficiently infect the cyst and ensure vertical transmission (Frydman et al., 2006). We hypothesized that if *Wolbachia* could not properly localize to the SSCN in *sim-bam-yfp;bam*$^{-}$ ovaries that this could result in its reduced titer. We looked at *Wolbachia* accumulation using an antibody to Hsp60 which has been shown to cross-react with *Wolbachia* (McGraw et al., 2002; Ferree et al., 2005; Serbus and Sullivan, 2007). We find that as in *mel-bam-yfp;bam*$^{-}$ flies, *Wolbachia* accumulates within the SSCN in *sim-bam-yfp;bam*$^{-}$ flies (Figure 4.6).

4.4 Discussion

The *bag of marbles* gene has diverged under rapid, adaptive evolution in *D. melanogaster* and *D. simulans* (Civetta et al., 2006; Bauer DuMont et al., 2007). In Chpater 3, we presented data suggesting that the selective force driving the adaptation of *bam* is acting in the female germline. We had hypothesized that the selective force could be the bacterial endosymbiont, *Wolbachia pipientis*, due to its maternal transmission and its ability to manipulate the production of the hosts that it infects (reviewed in Serbus et al., 2008; Werren et al., 2008). In this study, we find that *Wolbachia* interacts with *bam* by enhancing the fertility of *D. melanogaster bam* hypomorphs as well as *sim-bam-yfp;bam*$^{-}$ flies. Furthermore, we find the *D. melanogaster Wolbachia* cannot replicate to as high of levels in *sim-bam-yfp;bam*$^{-}$ ovaries.

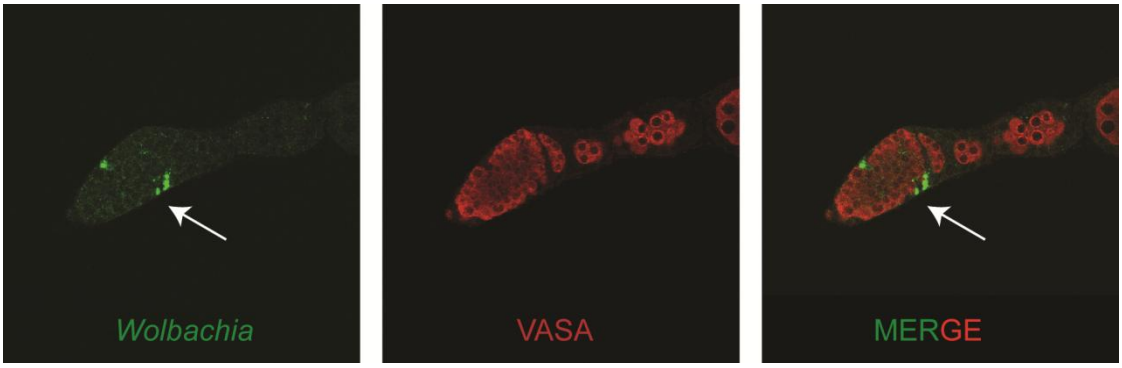


Figure 4.6: *Wolbachia* localizes to the SSCN in *sim-bam-yfp;bam⁻ flies.*
Ovaries from *sim-bam-yfp;bam⁻* flies are stained with antibodies to Vasa (red) and Hsp-60 (green) which recognizes *Wolbachia*. *Wolbachia* preferentially accumulate at the somatic stem cell niche (arrow) of the germarium.

4.4.1 *Wolbachia* and *bam*: direct or indirect?

Our results suggest that *Wolbachia* increases *bam* activity. This may occur either through a direct or indirect interaction with *bam*. If *Wolbachia* were acting directly on *bam*, it may do so by increasing *bam* RNA or protein levels. We measured *bam* RNA in *D. melanogaster bam* hypomorphs infected with *Wolbachia* relative to *bam* hypomorphs that had been cured of *Wolbachia* through tetracycline treatment. We find no evidence of increased *bam* RNA levels in *Wolbachia*-infected flies, suggesting that the increase in *bam* activity is not at the RNA level. It is possible that the presence of *Wolbachia* increases the amount of Bam protein present through reduced protein turnover or a difference in the rate of translation. We plan to use anti-Bam antibodies and Western blot analysis to test this in the future. Alternatively, *Wolbachia* may increase *bam* activity by allowing for more active Bam protein. A small amount of Bam protein is present in the GSC. Most of the protein is localized to the spectrosome, a small ER-like organelle, in the GSC where it is thought to be inactive in GSC differentiation (McKearin and Ohlstein, 1995). It is possible that *Wolbachia* interferes with efficient localization of Bam to the spectrosome, allowing more cytoplasmic Bam to participate in GSC differentiation.

It is also possible that *Wolbachia* is acting on regulators upstream or downstream of *bam*. While most Bam protein in the GSC is localized to the fusome, a small amount of Bam protein can be seen in the cytoplasm of GSCs (Jin et al., 2008). It has been hypothesized that there are proteins present within the GSC that inhibit any cytoplasmic Bam protein that is present (Xie, T. 2010 Drosophila Research Conference). It is possible that *Wolbachia* may inhibit this protein(s) from

antagonizing Bam in the GSC, thus allowing for more Bam protein to be available to stimulate GSC differentiation. The data presented in this study do not allow us to distinguish between direct or indirect interactions of *Wolbachia* and *bam.*

4.4.2 *Wolbachia, bam* and *Sex lethal*

Recent results indicate that *bam*'s function in GSC differentiation requires the function of *Sex lethal* (*Sxl*) (Chau et al., 2009). *Wolbachia* has also been shown to partially rescue the female sterility of *Sxl* mutants in *D. melanogaster*. This interaction was found to be allele-specific, suggesting it was unlikely for suppression to occur by a general increase in germline *Sxl* expression (Starr and Cline, 2002). Additionally, microarray studies on previtellogenic ovaries showed no significant increase in *Sxl* expression when infected with *Wolbachia* (Sun and Cline, 2009). *Sxl* is expressed in both GSCs and cystoblasts, while *bam* expression is repressed in GSCs and Bam protein is only present in cystoblasts and mitotically-active cysts. Therefore, it has been proposed that *Sxl* partners with newly-expressed Bam in cystoblasts to promote differentiation by antagonizing genes required to maintain GSCs (Chau et al., 2009). Based on consideration of the above data, we propose a model in which the increased fertility of *Wolbachia*-infected *Sxl* mutants is a result of increased *bam* activity driving the differentiation of GSCs, rather than a direct effect on *Sxl* activity. While forced expression of *bam* in *snf* mutants, which remove all *Sxl* from the germline, did not result in loss of GSCs, it would be interesting to determine if forced expression of *bam* in *Sxl* hypomorphs resulted in the differentiation of GSCs, the presence of which would support this model. However, until we know the molecular

nature of how *Wolbachia* increases the activity of *bam*, it will be difficult to test this hypothesis.

4.4.3 How might *Wolbachia* drive the adaptive evolution of *bam*?

We have shown an interaction between *bam* and *Wolbachia*. We have thus been interested in evaluating the hypothesis that *Wolbachia* has driven the adaptive evolution of *bam*. It is incredibly difficult to prove that a selective pressure drove the adaptive evolution of a gene when this event occurred in the past. Instead we propose models for how the interaction between *Wolbachia* and *bam* might have driven the adaptive evolution of *bam* and determine whether the predictions of these models are met in the data presented in this study.

We propose two models for how this interaction may drive the adaptive evolution of *bam*. In the first model, the initial introduction of the novel endosymbiont, *Wolbachia*, caused a misexpression of *bam* either at the RNA and/or protein level. The parasitic wasp, *Asobara tabida*, is a classic example of the novel pathogen model. The presence of *Wolbachia* is required for oogenesis to occur properly (Dedeine et al., 2001; Dedeine et al., 2005). It has been proposed that the initial *Wolbachia* infection resulted in suppression of normal host apoptosis that occurs during oocyte production. In response, the wasp has adapted by upregulating apoptosis. This response, while beneficial in the presence of *Wolbachia*, results in hyperactive apoptosis and oogenesis inhibition in its absence (Pannebakker et al., 2007). Additionally, *Wolbachia* infection alters the expression levels of numerous RNAs and proteins (Brennan et al., 2008; Xi et al., 2008; Kremer et al., 2009; Kremer

et al., 2012). Proper control of *bam* is critical as improper expression of *bam* in GSCs will cause precocious differentiation, while decreased *bam* expression can lead to an inability of germline cells to differentiate giving rise to ovarian tumors (McKearin and Spradling, 1990; Ohlstein and McKearin, 1997). If *bam* misexpression occurred in flies infected with *Wolbachia*, strong, directional selection would act on *bam* to restore its proper expression.

In this first model, we would predict that *bam* RNA and/or protein levels would be different in the presence of *Wolbachia*. We measured *bam* RNA levels between infected and uninfected *D. melanogaster bam* hypomorphs yet find no evidence of altered *bam* RNA levels in the presence of *Wolbachia*. In the future, we want to compare Bam protein levels in infected and uninfected *bam* hypomorphs as well to fully test this model. It is important to note that this model does not require a direct interaction on *bam* RNA or protein levels. *Wolbachia* infection could result in more active Bam protein by inhibiting a protein necessary to inhibit Bam function, for example. This interaction would still result in more Bam activity and could still result in directional selection acting on Bam to return it to a level that was beneficial for the host. We do not know much about potential inhibitors of Bam, so we cannot currently test any additional factors.

In our second model, we propose that *Wolbachia* is manipulating *bam* or utilizing *bam* expression or localization to "know" when to replicate or infect nearby cysts. This interaction could be deleterious for the host either by altering host reproduction or by allowing successful transmission of *Wolbachia*, allowing

Wolbachia to elicit other reproductive manipulations such as CI. This conflict between host and endosymbiont could lead to an evolutionary arms race.

This second model makes at least two predictions. The first is that both host and endosymbiont proteins involved in this interaction would adaptively evolve. In the case of *bam*, that expectation is met; it is adaptively evolving in both *D. melanogaster* and *D. simulans*. In the case of *Wolbachia*, while we do not know which genes are responsible for this interaction, the *Wolbachia* genomes of *D. melanogaster* strains and *D. simulans* strains differ dramatically in their content. There has been a rapid diversification of ankyrin repeat domain-containing genes between the two strains (Wu et al., 2004; Klasson et al., 2009), and ankyrin repeats are known to mediate protein-protein interactions (Caturegli et al., 2000). This diversification of ankyrin repeat-containing genes has been proposed to reflect diversification in the host molecules they target (Klasson et al., 2009). A second prediction is that each *Wolbachia* strain will have coevolved with its species-specific *bam* ortholog and that the success of the *Wolbachia* will be reduced in the presence of a heterospecific *bam* ortholog. We assayed *Wolbachia* titer and found that *D. melanogaster*-specific *Wolbachia* cannot accumulate as well in *sim-bam-yfp;bam*$^{-}$ ovaries. *sim-bam-yfp;bam*$^{-}$ ovaries show evidence of GSC loss, mitotic synchrony defects, improper regulation of cyst divisions. While we specifically used young flies to minimize the effects of different cell types, we cannot fully eliminate the possibility that differences in *w*Mel titer level could be due to differences in cellular types present. However, in the future, I plan to take a cytological approach comparing *Wolbachia* titer levels in similarly-staged egg chambers (described in Serbus and

Sullivan, 2007; Serbus et al., 2011) between the two transgenic lines to alleviate worries that differences in *Wolbachia* titer are due to cellular defects and not functional divergence.

4.4.2 The effects of *Wolbachia* infection on insect biology

Using PCR-based assays, it has been estimated by Clark et al. (2005) that approximately 30% of *D. melanogaster* stocks in the Bloomington Drosophila Stock Center are infected with *Wolbachia*. This study did not survey for infection with the other known *D. melanogaster* endosymbiont, *Spiroplasma*. Therefore, this estimate of endosymbiont infections is conservative. Additionally, surveys have found *Wolbachia* and Spiroplasma to be prevalent across Drosophila species (Mateos et al., 2006; Watts et al., 2009). We concur with others that these findings should not result in a wide-spread antibiotic treatment to cure stocks of *Wolbachia*-infection (Clark et al., 2005). On the contrary, our results emphasize the importance of understanding developmental biology and evolution in the presence of bacterial endosymbionts, as they can drastically alter our interpretation of the gene-by-gene and gene-by-environment interactions we observe and have a dramatic impact on the evolution of host genes.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

In the Aquadro lab, *bam* was originally surveyed for nucleotide variation in *D. melanogaster* and *D. simulans* because it lies in an area of relatively high recombination and given its critical function in GSC differentiation, it was assumed *bam* would evolve under strong, functional constraint. Surprisingly, *bam* was found to be rapidly diverging under adaptive evolution in both *D. melanogaster* and *D. simulans* (Civetta et al., 2006; Bauer DuMont et al., 2007). Since *bam* is expressed much earlier in gametogenesis, it is unlikely that *bam* is subject to selective pressures such as sexual conflict and sexual selection, pressures that are often invoked to explain the adaptive evolution of reproductive genes. What then, is driving the adaptive evolution of *bam*? And what functions of *bam* are being altered?

My thesis was designed to generate molecular-genetic tools and apply them to explore the functional consequences of *bam*'s sequence divergence. Our hypothesis was that the functions of *bam* that have diverged under adaptive evolution would be the functions for which we would not see full rescue when using *sim-bam-yfp*. Once we knew the functional consequences of the divergence of *bam*, we could utilize that information to generate and ultimately test hypotheses about the selective force(s) that have driven the adaptive evolution of *bam*. While this methodology proved to be successful in many ways, we found a number of difficulties that we had not

anticipated. I believe this experimental approach can continue to be successful in the future; however, I have a number of suggestions that I think will improve the use of transgenes to determine the functional divergence of genes. I also provide suggestions regarding future work on *bam*.

5.1 *mel-bam-yfp* under-expression

The causes of the reduced expression of *mel-bam-yfp* relative to the endogenous locus remain unclear. If further experiments are to be done to assay functional divergence of *bam*, it would be essential to make new transgenes. To do so, we first need to determine if *mel-bam-yfp* is expressing so low because critical portions of the regulatory sequences are missing that are necessary for proper *bam* expression. One way to address this question would be to order BAC clones that include *bam* from BACPAC Resources (http://bacpac.chori.org/home.htm). This collection includes two 20 kb BACs (CH322-1H3 and CH322-16B9) and one 80 kb BAC (CH321-71O5) that span *bam*. If it is found that the larger BAC containing *bam* provides both full function and expression rescue, the two smaller BACs may help to provide narrower boundaries on regulatory regions as one includes more 5' regulatory sequence and the other includes more 3' regulatory sequence spanning *bam*.

5.2 *bam* and *Wolbachia*

While I found evidence for an interaction between *bam* and *Wolbachia*, we still do not know if this is due to *Wolbachia* directly interacting with *bam* or genes that are functionally upstream or downstream of *bam*. Unfortunately, we currently do not

know about other genes that antagonize *bam* activity. If we did, we could test for an interaction between *Wolbachia* and those genes. If an interaction existed, we could determine if this was due to altered RNA and protein levels in the presence of *Wolbachia*. One experiment I did perform was to determine if I could generate a hypomorphic phenotype by crossing together any two mutant alleles of *bgcn*. All 6 alleles I tried produced the null phenotype of completely tumorous ovaries. Given that rescue with both *bam* and *Sxl* was only seen with hypomorphs, I felt it was not worth testing if *Wolbachia* infection rescues a *bgcn* null mutation. We do know that *bam* functions to antagonize *nanos* and *pumilio* (Li et al., 2009). One could cross *Wolbachia* into *nanos* or *pumilio* hypomorphs and ask if there is rescue of those phenotypes. If rescue occurred, one could ask if *nanos* or *pumilio* RNA or protein levels are altered to ask if the increase in *bam* activity we see was actually due to *Wolbachia* interacting with these downstream components. Hypomorphs of *nanos* have been reported (Li et al., 2009), so this experiment is relatively easy to do. While these suggestions provide follow-up experiments directly related to my dissertation work, my dissertation has also provided us with valuable insight of how one should perform transgenic assays in the future.

5.3 ΦC31-mediated integration

ΦC31-mediated site-specific transgene integration is a huge advancement in transgenesis technology over the original random insertions of P-element-mediated transformation vectors. It allows for the integration of any *attB*-containing transgene into the same *attP* docking site within the genome, thus removing variation due to

position effects. The appeal of this system was that it would require much less work in generating transformants. The system itself produced transformants at a much higher frequency (Groth et al., 2004), and inserting transgenes into the same position of the genome abrogates the need to compare multiple independently-derived transgenic lines.

I used the ΦC31 transgenesis system for my experiments described in Chapters 3 and 4 and have found it to be extremely useful for paired transgene comparisons. However, I describe two necessary controls that my experiences, and those of the Barbash lab, indicate are required to ensure proper interpretation of experiments using the ΦC31-integration system.

First, use positive controls when determining transgene integration by PCR-based assays. Transgene integration is confirmed by inability to amplify the *attP* amplicon and the gain of the ability to amplify the *attL* and *attR* amplicons (Venken et al., 2006). Using the conditions described in Venken *et al*. (2006), I confirmed transgene integration and the apparent lack of an unoccupied *attP* site. However, I later determined that the attP16 stock that was injected contained two *attP* sites. Even though I used published protocols, the *attP* PCR assay was not optimized and thus gave no product indicating the lack of transgene integration into the second site. I was not able to realize this because I had no positive control for an unintegrated *attP* site for the PCR. We now have lines in the lab that can serve as positive controls for all three PCR assays (*mel-bam-yfp* line 29-1 or *sim-bam-yfp* line 1-1 for *attL*,*attR*, and *attP* amplification). Additionally, we have developed PCR assays that are specific to each *attP* docking site to ensure that transgenes are integrated into the correct sites.

Second, while the ΦC31-integration system may eliminate variation between transgenes within the same docking site, it has been shown that different docking sites still cause dramatic differences in transgene expression (Markstein et al., 2008). In the Barbash lab, we have come across a situation where comparisons of paired transgenes gave opposite results depending on the *attP* site into which they were integrated. Therefore, it is necessary to compare transgene pairs docked at different *attP* sites to confirm that the results obtained are similar. Furthermore, comparisons made across different insertion sites provide more power in the instance where the differences may be small. For example, in Chapter 4, I presented data that *Wolbachia* infection resulted in a small, but statistically significant increase in the fertility of young *sim-bam-yfp*;*bam*$^{-}$ flies. By having performed that experiment using transgenes in a different insertion site, we had two independent sources of data yielding similar results, which provided confidence that these data are, in fact, biologically meaningful.

5.4 Transgene Rescue Assays

In Chapter 3, I described a transgene rescue assay of *mel-bam-yfp* or *sim-bam-yfp* to rescue the female and male sterility of a *D. melanogaster bam* mutant. Based on my experiences as well as those of other members in the Barbash Lab, I have three suggestions for future work using transgene rescue assays.

First, when designing constructs it is important to determine whether published transgene rescue experiments were performed in a qualitative or quantitative manner. In the case of *bam*, previous rescue assays were done qualitatively. Transgenes were simply assayed as to whether they rescued male and female sterility, not whether they

quantitatively rescued sterility (McKearin and Spradling, 1990; Chen and McKearin, 2003b). While for most uses this measure would not be a problem, in our assays, we expected and needed full rescue of the *mel-bam-yfp* transgene. For future rescue assays, transgenes should be assayed in a quantitative manner prior to making additional transgenes and/or modifications to transgenes.

Second, in the transgene rescue assay, we developed two transgenes. The first was a control that was the *D. melanogaster bam* ORF tagged with YFP and driven by the *D. melanogaster bam* regulatory region. The second transgene was the *D. simulans bam* ORF tagged with YFP and driven by the *D. melanogaster* regulatory region; this was designed in an effort to attribute any differences we saw to coding-sequence divergence. When we developed these transgenes, we considered making additional controls such as *D. simulans bam* ORF with the *D. simulans* regulatory region or untagged versions of each transgene, but ultimately decided against these controls. In retrospect, these control transgenes would have been extremely valuable. What if there is an incompatibility between the *D. melanogaster* regulatory region and the *D. simulans bam* ORF that is causing the lack of rescue of *sim-bam-yfp* rather than coding sequence divergence? Or, what if the YFP tag inhibits activity of *sim-bam-yfp*, but not *mel-bam-yfp*? While one can try to hand-wave these alternatives away, the bottom line is that they are possible alternative interpretations of these data. Having these controls allows for one to easily eliminate these explanations for the differences we see in transgenic *bam* comparisons.

Third, multiple measures must be used to assay transgene rescue. Measures of rescue should include both the most stringent functional measures as well as rescue of

expression level and pattern. In our experiments, *mel-bam-yfp* rescued both female and male sterility under normal fertility conditions. It was only under sperm exhaustion conditions where we saw lack of rescue in males. Furthermore, it was only when we used a *D. melanogaster bam* heterozygote as a control in our experiments that it became clear that the problem was with *mel-bam-yfp* and not *sim-bam-yfp*. While *mel-bam-yfp* rescued female sterility, its expression was dramatically reduced at the RNA level, thus limiting our ability to interpret our data. In experiments performed by Shamoni Maheshwari in the lab using *Lhr* transgenes, she found that while *mel-Lhr-HA* was expressed at wildtype levels at the RNA and protein levels, it did not fully perform like the endogenous locus (Maheshwari and Barbash, 2012). These two examples emphasize the need to use both comparisons to heterozygous controls and to measure expression to assay full rescue.

Insco *et al*. (2009) showed that alterations in *bam* expression change the number of divisions a germline cyst undergoes in spermatogenesis. While the number of cyst divisions is highly conserved among females across Drosophila species, this is not true for males. It has been hypothesized that *bam* may play a role in regulating these differences among male Drosophila species (Fuller, 1998). Therefore, when I started this project, it seemed very likely that selective pressures in the male germline would be driving the adaptive evolution of *bam*. Instead, our results suggested that selective pressures in the female germline were driving the adaptive evolution of *bam*. While I expected to learn about this interspecific divergence of *bam*, I was surprised to also gain insight into the intraspecific functions of *bam* as well. Our qRT-PCR experiments measuring *bam* expression in *mel-bam-yfp*;*bam*$^{-}$ flies show that the

regulatory region that is necessary to drive proper *bam* expression has been inaccurately defined. Furthermore, comparing *bam* expression between *mel-bam-yfp*;*bam*$^{-}$ and *sim-bam-yfp*;*bam*$^{-}$ flies showed that while *bam* is under-expressed at the RNA level, there is little difference in protein levels. This could suggest a feedback mechanism acting on *bam* protein levels, a project that could be explored in the future. In my experiments, I also found an interaction between *D. melanogaster bam* and *Wolbachia*. While we currently do not know the nature of this interaction, future work might elucidate how this interaction affects both *bam* and *Wolbachia* in a wildtype fly.

Reference List

Aagaard, J.E., Vacquier, V.D., MacCoss, M.J., and Swanson, W.J. (2010). ZP domain proteins in the abalone egg coat include a paralog of VERL under positive selection that binds lysin and 18-kDa sperm proteins Mol. Biol. Evol. *27,* 193-203.

Aguade, M., Miyashita, N., and Langley, C.H. (1992). Polymorphism and divergence in the Mst26A male accessory gland gene region in Drosophila Genetics *132,* 755-770.

Anderson, J.A., Gilliland, W.D., and Langley, C.H. (2009). Molecular population genetics and evolution of Drosophila meiosis genes Genetics *181,* 177-185.

Andres, J.A., Maroja, L.S., Bogdanowicz, S.M., Swanson, W.J., and Harrison, R.G. (2006). Molecular evolution of seminal proteins in field crickets Mol. Biol. Evol. *23,* 1574-1584.

Aquadro, C.F., Lado, K.M., and Noon, W.A. (1988). The rosy region of *Drosophila melanogaster* and *Drosophila simulans*. I. Contrasting levels of naturally occurring DNA restriction map variation and divergence. Genetics *119,* 875-888.

Aravin, A.A., Hannon, G.J., and Brennecke, J. (2007). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race Science *318,* 761-764.

Aruna, S., Flores, H.A., and Barbash, D.A. (2009). Reduced fertility of *Drosophila melanogaster hybrid male rescue* (*Hmr*) mutant females is partially complemented by *Hmr* orthologs from sibling species. Genetics *181,* 1437-1450.

Bauer DuMont, V.L., Flores, H.A., Wright, M.H., and Aquadro, C.F. (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.

Begun, D.J., and Aquadro, C.F. (1994). Evolutionary inferences from DNA variation at the 6-phosphogluconate dehydrogenase locus in natural populations of drosophila: selection and geographic differentiation Genetics *136,* 155-171.

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

Begun, D.J., Holloway, A.K., Stevens, K., Hillier, L.W., Poh, Y.P., Hahn, M.W., Nista, P.M., Jones, C.D., Kern, A.D., Dewey, C.N., *et al.* (2007). Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. PLoS Biol. *5,* e310.

Begun, D.J., Whitley, P., Todd, B.L., Waldrip-Dail, H.M., and Clark, A.G. (2000). Molecular population genetics of male accessory gland proteins in Drosophila Genetics *156,* 1879-1888.

Boyle, L., O'Neill, S.L., Robertson, H.M., and Karr, T.L. (1993). Interspecific and intraspecific horizontal transfer of *Wolbachia* in Drosophila. Science *260,* 1796-1799.

Brennan, L.J., Keddie, B.A., Braig, H.R., and Harris, H.L. (2008). The endosymbiont *Wolbachia pipientis* induces the expression of host antioxidant proteins in an *Aedes albopictus* cell line. PLoS One *3,* e2083.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing Science *322,* 1387-1392.

Bustamante, C.D., Fledel-Alon, A., Williamson, S., Nielsen, R., Hubisz, M.T., Glanowski, S., Tanenbaum, D.M., White, T.J., Sninsky, J.J., Hernandez, R.D., *et al.* (2005). Natural selection on protein-coding genes in the human genome Nature *437,* 1153-1157.

Caturegli, P., Asanovich, K.M., Walls, J.J., Bakken, J.S., Madigan, J.E., Popov, V.L., and Dumler, J.S. (2000). ankA: an Ehrlichia phagocytophila group gene encoding a cytoplasmic protein antigen with ankyrin repeats Infect. Immun. *68,* 5277-5283.

Chau, J., Kulnane, L.S., and Salz, H.K. (2009). *Sex-lethal* facilitates the transition from germline stem cell to committed daughter cell in the Drosophila ovary. Genetics *182,* 121-132.

Chen, D., and McKearin, D. (2005). Gene circuitry controlling a stem cell niche. Curr. Biol. *15,* 179-184.

Chen, D., and McKearin, D. (2003a). Dpp signaling silences *bam* transcription directly to establish asymmetric divisions of germline stem cells. Curr. Biol. *13,* 1786-1791.

Chen, D., and McKearin, D.M. (2003b). A discrete transcriptional silencer in the *bam* gene determines asymmetric division of the Drosophila germline stem cell. Development *130,* 1159-1170.

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

Civetta, A., and Singh, R.S. (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.

Clark, K.A., and McKearin, D.M. (1996). The Drosophila *stonewall* gene encodes a putative transcription factor essential for germ cell development. Development *122,* 937-950.

Clark, M.E., Anderson, C.L., Cande, J., and Karr, T.L. (2005). Widespread prevalence of Wolbachia in laboratory stocks and the implications for Drosophila research. Genetics *170,* 1667-1675.

Clark, N.L., Aagaard, J.E., and Swanson, W.J. (2006). Evolution of reproductive proteins from animals and plants Reproduction *131,* 11-22.

Clark, N.L., and Swanson, W.J. (2005). Pervasive adaptive evolution in primate seminal proteins PLoS Genet. *1,* e35.

Cox, D.N., 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.

Cutter, A.D., and Ward, S. (2005). Sexual and temporal dynamics of molecular evolution in *C. elegans* development. Mol. Biol. Evol. *22,* 178-188.

de Cuevas, M., Lee, J.K., and Spradling, A.C. (1996). *alpha-spectrin* is required for germline cell division and differentiation in the Drosophila ovary. Development *122,* 3959-3968.

de Cuevas, M., and Spradling, A.C. (1998). Morphogenesis of the Drosophila fusome and its implications for oocyte specification Development *125,* 2781-2789.

Dean, M.D., and Ballard, J.W. (2004). Linking phylogenetics with population genetics to reconstruct the geographic origin of a species Mol. Phylogenet. Evol. *32,* 998-1009.

Dean, M.D., Clark, N.L., Findlay, G.D., Karn, R.C., Yi, X., Swanson, W.J., MacCoss, M.J., and Nachman, M.W. (2009). Proteomics and comparative genomic investigations reveal heterogeneity in evolutionary rate of male reproductive proteins in mice (*Mus domesticus*). Mol. Biol. Evol. *26,* 1733-1743.

Decotto, E., and Spradling, A.C. (2005). The Drosophila ovarian and testis stem cell niches: similar somatic stem cells and signals Dev. Cell. *9,* 501-510.

Dedeine, F., Bouletreau, M., and Vavre, F. (2005). *Wolbachia* requirement for oogenesis: occurrence within the genus *Asobara* (Hymenoptera, Braconidae) and evidence for intraspecific variation in *A. tabida*. Heredity (Edinb) *95,* 394-400.

Dedeine, F., Vavre, F., Fleury, F., Loppin, B., Hochberg, M.E., and Bouletreau, M. (2001). Removing symbiotic *Wolbachia* bacteria specifically inhibits oogenesis in a parasitic wasp. Proc. Natl. Acad. Sci. U. S. A. *98*, 6247-6252.

Drosophila 12 Genomes Consortium, Clark, A.G., Eisen, M.B., Smith, D.R., Bergman, C.M., Oliver, B., Markow, T.A., Kaufman, T.C., Kellis, M., Gelbart, W., *et al.* (2007). Evolution of genes and genomes on the Drosophila phylogeny Nature *450*, 203-218.

DuMont, V.B., Fay, J.C., Calabrese, P.P., and Aquadro, C.F. (2004). DNA variability and divergence at the *notch* locus in *Drosophila melanogaster* and *D. simulans*: a case of accelerated synonymous site divergence. Genetics *167*, 171-185.

Fast, E.M., Toomey, M.E., Panaram, K., Desjardins, D., Kolaczyk, E.D., and Frydman, H.M. (2011). *Wolbachia* enhance Drosophila stem cell proliferation and target the germline stem cell niche. Science *334*, 990-992.

Fay, J.C., and Wu, C.I. (2000). Hitchhiking under positive Darwinian selection Genetics *155*, 1405-1413.

Ferree, P.M., Frydman, H.M., Li, J.M., Cao, J., Wieschaus, E., and Sullivan, W. (2005). *Wolbachia* utilizes host microtubules and Dynein for anterior localization in the Drosophila oocyte. PLoS Pathog. *1*, e14.

Forbes, A., and Lehmann, R. (1998). Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells Development *125*, 679-690.

Frand, A.R., Russel, S., and Ruvkun, G. (2005). Functional genomic analysis of *C. elegans* molting. PLoS Biol. *3*, e312.

Fry, A.J., Palmer, M.R., and Rand, D.M. (2004). Variable fitness effects of *Wolbachia* infection in *Drosophila melanogaster*. Heredity (Edinb) *93*, 379-389.

Frydman, H.M., Li, J.M., Robson, D.N., and Wieschaus, E. (2006). Somatic stem cell niche tropism in *Wolbachia*. Nature *441*, 509-512.

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

Fuller, M.T., and Spradling, A.C. (2007). Male and female Drosophila germline stem cells: two versions of immortality Science *316*, 402-404.

Gilboa, L., Forbes, A., Tazuke, S.I., Fuller, M.T., and Lehmann, R. (2003). Germ line stem cell differentiation in Drosophila requires gap junctions and proceeds via an intermediate state Development *130,* 6625-6634.

Gilboa, L., and Lehmann, R. (2004). How different is Venus from Mars? The genetics of germ-line stem cells in Drosophila females and males Development *131,* 4895-4905.

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.

Good, J.M., and Nachman, M.W. (2005). Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis Mol. Biol. Evol. *22,* 1044-1052.

Groth, A.C., Fish, M., Nusse, R., and Calos, M.P. (2004). Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31 Genetics *166,* 1775-1782.

Haddrill, P.R., Halligan, D.L., Tomaras, D., and Charlesworth, B. (2007). Reduced efficacy of selection in regions of the Drosophila genome that lack crossing over Genome Biol. *8,* R18.

Haerty, W., Jagadeeshan, S., Kulathinal, R.J., Wong, A., Ravi Ram, K., Sirot, L.K., Levesque, L., Artieri, C.G., Wolfner, M.F., Civetta, A., and Singh, R.S. (2007). Evolution in the fast lane: rapidly evolving sex-related genes in Drosophila Genetics *177,* 1321-1335.

Harris, R.E., Pargett, M., Sutcliffe, C., Umulis, D., and Ashe, H.L. (2011). Brat promotes stem cell differentiation via control of a bistable switch that restricts BMP signaling Dev. Cell. *20,* 72-83.

Hawkins, N.C., Thorpe, J., and Schupbach, T. (1996). *encore*, a gene required for the regulation of germ line mitosis and oocyte differentiation during Drosophila oogenesis. Development *122,* 281-290.

Hedges, L.M., Brownlie, J.C., O'Neill, S.L., and Johnson, K.N. (2008). *Wolbachia* and virus protection in insects. Science *322,* 702.

Hey, J., and Kliman, R.M. (2002). Interactions between natural selection, recombination and gene density in the genes of Drosophila Genetics *160,* 595-608.

Hoekstra, H.E., Hirschmann, R.J., Bundey, R.A., Insel, P.A., and Crossland, J.P. (2006). A single amino acid mutation contributes to adaptive beach mouse color pattern Science *313,* 101-104.

Hollocher, H., Agopian, K., Waterbury, J., O'Neill, R.W., and Davis, A.W. (2000). Characterization of defects in adult germline development and oogenesis of sterile and rescued female hybrids in crosses between *Drosophila simulans* and *Drosophila melanogaster*. J. Exp. Zool. *288,* 205-218.

Hornett, E.A., Charlat, S., Duplouy, A.M., Davies, N., Roderick, G.K., Wedell, N., and Hurst, G.D. (2006). Evolution of male-killer suppression in a natural population PLoS Biol. *4,* e283.

Hussain, M., Frentiu, F.D., Moreira, L.A., O'Neill, S.L., and Asgari, S. (2011). *Wolbachia* uses host microRNAs to manipulate host gene expression and facilitate colonization of the dengue vector *Aedes aegypti*. Proc. Natl. Acad. Sci. U. S. A. *108,* 9250-9255.

Insco, M.L., Leon, A., Tam, C.H., McKearin, D.M., and Fuller, M.T. (2009). Accumulation of a differentiation regulator specifies transit amplifying division number in an adult stem cell lineage Proc. Natl. Acad. Sci. U. S. A. *106,* 22311-22316.

Jensen, J.D., Wong, A., and Aquadro, C.F. (2007). Approaches for identifying targets of positive selection Trends Genet. *23,* 568-577.

Jin, Z., Kirilly, D., Weng, C., Kawase, E., Song, X., Smith, S., Schwartz, J., and Xie, T. (2008). Differentiation-defective stem cells outcompete normal stem cells for niche occupancy in the Drosophila ovary Cell. Stem Cell. *2,* 39-49.

Kadyrova, L.Y., Habara, Y., Lee, T.H., and Wharton, R.P. (2007). Translational control of maternal *Cyclin B* mRNA by Nanos in the Drosophila germline. Development *134,* 1519-1527.

Kawase, E., Wong, M.D., Ding, B.C., and Xie, T. (2004). Gbb/Bmp signaling is essential for maintaining germline stem cells and for repressing *bam* transcription in the Drosophila testis. Development *131,* 1365-1375.

Kelleher, E.S., and Markow, T.A. (2009). Duplication, selection and gene conversion in a *Drosophila mojavensis* female reproductive protein family. Genetics *181,* 1451-1465.

Kelleher, E.S., Swanson, W.J., and Markow, T.A. (2007). Gene duplication and adaptive evolution of digestive proteases in *Drosophila arizonae* female reproductive tracts. PLoS Genet. *3,* e148.

King, F.J., and Lin, H. (1999). Somatic signaling mediated by fs(1)Yb is essential for germline stem cell maintenance during Drosophila oogenesis Development *126*, 1833-1844.

King, F.J., Szakmary, A., Cox, D.N., and Lin, H. (2001). *Yb* modulates the divisions of both germline and somatic stem cells through *piwi*- and *hh*-mediated mechanisms in the Drosophila ovary. Mol. Cell *7*, 497-508.

Kirilly, D., and Xie, T. (2007). The Drosophila ovary: an active stem cell community Cell Res. *17*, 15-25.

Klasson, L., Westberg, J., Sapountzis, P., Naslund, K., Lutnaes, Y., Darby, A.C., Veneti, Z., Chen, L., Braig, H.R., Garrett, R., Bourtzis, K., and Andersson, S.G. (2009). The mosaic genome structure of the *Wolbachia w*Ri strain infecting *Drosophila simulans*. Proc. Natl. Acad. Sci. U. S. A. *106*, 5725-5730.

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J.S., Zhang, F., Schultz, N., Koppetsch, B.S., Nowosielska, A., *et al.* (2009). The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters Cell *138*, 1137-1149.

Kolaczkowski, B., Hupalo, D.N., and Kern, A.D. (2011). Recurrent adaptation in RNA interference genes across the Drosophila phylogeny Mol. Biol. Evol. *28*, 1033-1042.

Kosiol, C., Vinar, T., da Fonseca, R.R., Hubisz, M.J., Bustamante, C.D., Nielsen, R., and Siepel, A. (2008). Patterns of positive selection in six Mammalian genomes PLoS Genet. *4*, e1000144.

Kreitman, M. (1983). Nucleotide polymorphism at the *alcohol dehydrogenase* locus of *Drosophila melanogaster*. Nature *304*, 412-417.

Kremer, N., Charif, D., Henri, H., Bataille, M., Prevost, G., Kraaijeveld, K., and Vavre, F. (2009). A new case of *Wolbachia* dependence in the genus *Asobara*: evidence for parthenogenesis induction in *Asobara japonica*. Heredity (Edinb) *103*, 248-256.

Kremer, N., Charif, D., Henri, H., Gavory, F., Wincker, P., Mavingui, P., and Vavre, F. (2012). Influence of *Wolbachia* on host gene expression in an obligatory symbiosis. BMC Microbiol. *12 Suppl 1*, S7.

Larracuente, A.M., Sackton, T.B., Greenberg, A.J., Wong, A., Singh, N.D., Sturgill, D., Zhang, Y., Oliver, B., and Clark, A.G. (2008). Evolution of protein-coding genes in *Drosophila*. Trends Genet. *24*, 114-123.

Lavoie, C.A., Ohlstein, B., and McKearin, D.M. (1999). Localization and function of Bam protein require the *benign gonial cell neoplasm* gene product. Dev. Biol. *212,* 405-413.

Lawniczak, M.K., and Begun, D.J. (2007). Molecular population genetics of female-expressed mating-induced serine proteases in *Drosophila melanogaster*. Mol. Biol. Evol. *24,* 1944-1951.

Leon, A., and McKearin, D. (1999). Identification of TER94, an AAA ATPase protein, as a Bam-dependent component of the *Drosophila* fusome. Mol. Biol. Cell *10,* 3825-3834.

Li, V.C., Davis, J.C., Lenkov, K., Bolival, B., Fuller, M.T., and Petrov, D.A. (2009). Molecular evolution of the testis TAFs of Drosophila Mol. Biol. Evol. *26,* 1103-1116.

Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C.C., and Kain, S.R. (1998). Generation of destabilized green fluorescent protein as a transcription reporter J. Biol. Chem. *273,* 34970-34975.

Li, Y., Maines, J.Z., Tastan, O.Y., McKearin, D.M., and Buszczak, M. (2012a). Mei-P26 regulates the maintenance of ovarian germline stem cells by promoting BMP signaling Development *139,* 1547-1556.

Li, Y., Maines, J.Z., Tastan, O.Y., McKearin, D.M., and Buszczak, M. (2012b). Mei-P26 regulates the maintenance of ovarian germline stem cells by promoting BMP signaling Development *139,* 1547-1556.

Li, Y., Minor, N.T., Park, J.K., McKearin, D.M., and Maines, J.Z. (2009). Bam and Bgcn antagonize *Nanos*-dependent germ-line stem cell maintenance. Proc. Natl. Acad. Sci. U. S. A. *106,* 9304-9309.

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data Bioinformatics *25,* 1451-1452.

Lighthouse, D.V., Buszczak, M., and Spradling, A.C. (2008). New components of the *Drosophila* fusome suggest it plays novel roles in signaling and transport. Dev. Biol. *317,* 59-71.

Lilly, M.A., de Cuevas, M., and Spradling, A.C. (2000). Cyclin A associates with the fusome during germline cyst formation in the *Drosophila* ovary. Dev. Biol. *218,* 53-63.

Lin, H. (1998). The self-renewing mechanism of stem cells in the germline Curr. Opin. Cell Biol. *10,* 687-693.

Lin, H., and Spradling, A.C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary Development *124*, 2463-2476.

MacCallum, C., and Hill, E. (2006). Being positive about selection PLoS Biol. *4*, e87.

Macpherson, J.M., Sella, G., Davis, J.C., and Petrov, D.A. (2007). Genomewide spatial correspondence between nonsynonymous divergence and neutral polymorphism reveals extensive adaptation in Drosophila Genetics *177*, 2083-2099.

Maheshwari, S., and Barbash, D.A. (2012). Cis-by-Trans Regulatory Divergence Causes the Asymmetric Lethal Effects of an Ancestral Hybrid Incompatibility Gene PLoS Genet. *8*, e1002597.

Maines, J.Z., Park, J.K., Williams, M., and McKearin, D.M. (2007). Stonewalling Drosophila stem cell differentiation by epigenetic controls. Development *134*, 1471-1479.

Markstein, M., Pitsouli, C., Villalta, C., Celniker, S.E., and Perrimon, N. (2008). Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes Nat. Genet. *40*, 476-483.

Mateos, M., Castrezana, S.J., Nankivell, B.J., Estes, A.M., Markow, T.A., and Moran, N.A. (2006). Heritable endosymbionts of *Drosophila*. Genetics *174*, 363-376.

McDonald, J.H., and Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. Nature *351*, 652-654.

McGraw, E.A., Merritt, D.J., Droller, J.N., and O'Neill, S.L. (2002). *Wolbachia* density and virulence attenuation after transfer into a novel host. Proc. Natl. Acad. Sci. U. S. A. *99*, 2918-2923.

McGraw, E.A., Merritt, D.J., Droller, J.N., and O'Neill, S.L. (2001). *Wolbachia*-mediated sperm modification is dependent on the host genotype in *Drosophila*. Proc. Biol. Sci. *268*, 2565-2570.

McKearin, D., and Ohlstein, B. (1995). A role for the *Drosophila* Bag-of-marbles protein in the differentiation of cystoblasts from germline stem cells. Development *121*, 2937-2947.

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

Morel, N., and Massoulie, J. (2000). Comparative expression of homologous proteins. A novel mode of transcriptional regulation by the coding sequence folding compatibility of chimeras J. Biol. Chem. *275,* 7304-7312.

Moriyama, E.N., and Powell, J.R. (1996). Intraspecific nuclear DNA variation in *Drosophila*. Mol. Biol. Evol. *13,* 261-277.

Nachman, M.W., Hoekstra, H.E., and D'Agostino, S.L. (2003). The genetic basis of adaptive melanism in pocket mice Proc. Natl. Acad. Sci. U. S. A. *100,* 5268-5273.

Neumuller, R.A., Betschinger, J., Fischer, A., Bushati, N., Poernbacher, I., Mechtler, K., Cohen, S.M., and Knoblich, J.A. (2008). Mei-P26 regulates microRNAs and cell growth in the *Drosophila* ovarian stem cell lineage. Nature *454*, 241-245.

Neznanov, N., Umezawa, A., and Oshima, R.G. (1997). A regulatory element within a coding exon modulates keratin 18 gene expression in transgenic mice J. Biol. Chem. *272,* 27549-27557.

Obbard, D.J., Gordon, K.H., Buck, A.H., and Jiggins, F.M. (2009). The evolution of RNAi as a defence against viruses and transposable elements Philos. Trans. R. Soc. Lond. B. Biol. Sci. *364,* 99-115.

Ohlstein, B., Lavoie, C.A., Vef, O., Gateff, E., and McKearin, D.M. (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.

Ohlstein, B., and McKearin, D. (1997). Ectopic expression of the *Drosophila* Bam protein eliminates oogenic germline stem cells. Development *124,* 3651-3662.

Olivieri, D., Sykora, M.M., Sachidanandam, R., Mechtler, K., and Brennecke, J. (2010). An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila EMBO J. *29,* 3301-3317.

Ometto, L., Glinka, S., De Lorenzo, D., and Stephan, W. (2005). Inferring the effects of demography and selection on *Drosophila melanogaster* populations from a chromosome-wide scan of DNA variation. Mol. Biol. Evol. *22,* 2119-2130.

Panhuis, T.M., Clark, N.L., and Swanson, W.J. (2006). Rapid evolution of reproductive proteins in abalone and *Drosophila*. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *361,* 261-268.

Panhuis, T.M., and Swanson, W.J. (2006). Molecular evolution and population genetic analysis of candidate female reproductive genes in *Drosophila*. Genetics *173,* 2039-2047.

Pannebakker, B.A., Loppin, B., Elemans, C.P., Humblot, L., and Vavre, F. (2007). Parasitic inhibition of cell death facilitates symbiosis Proc. Natl. Acad. Sci. U. S. A. *104,* 213-215.

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

Poinsot, D., Bourtzis, K., Markakis, G., Savakis, C., and Mercot, H. (1998). *Wolbachia* transfer from *Drosophila melanogaster* into *D. simulans*: Host effect and cytoplasmic incompatibility relationships. Genetics *150,* 227-237.

Pool, J.E., and Aquadro, C.F. (2006). History and structure of sub-Saharan populations of *Drosophila melanogaster*. Genetics *174,* 915-929.

Presgraves, D.C. (2007). Does genetic conflict drive rapid molecular evolution of nuclear transport genes in *Drosophila*? Bioessays *29,* 386-391.

Prokupek, A., Hoffmann, F., Eyun, S.I., Moriyama, E., Zhou, M., and Harshman, L. (2008). An evolutionary expressed sequence tag analysis of *Drosophila* spermatheca genes. Evolution *62,* 2936-2947.

Rangan, P., Malone, C.D., Navarro, C., Newbold, S.P., Hayes, P.S., Sachidanandam, R., Hannon, G.J., and Lehmann, R. (2011). piRNA production requires heterochromatin formation in *Drosophila*. Curr. Biol. *21,* 1373-1379.

Rogers, S., Wells, R., and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis Science *234,* 364-368.

Saito, K., Ishizu, H., Komai, M., Kotani, H., Kawamura, Y., Nishida, K.M., Siomi, H., and Siomi, M.C. (2010). Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in *Drosophila*. Genes Dev. *24,* 2493-2498.

Schmidt, P.S., and Paaby, A.B. (2008). Reproductive diapause and life-history clines in North American populations of *Drosophila melanogaster*. Evolution *62,* 1204-1215.

Senti, K.A., and Brennecke, J. (2010). The piRNA pathway: a fly's perspective on the guardian of the genome Trends Genet. *26,* 499-509.

Serbus, L.R., Casper-Lindley, C., Landmann, F., and Sullivan, W. (2008). The genetics and cell biology of *Wolbachia*-host interactions. Annu. Rev. Genet. *42,* 683-707.

Serbus, L.R., Ferreccio, A., Zhukova, M., McMorris, C.L., Kiseleva, E., and Sullivan, W. (2011). A feedback loop between *Wolbachia* and the Drosophila *gurken* mRNP complex influences *Wolbachia* titer. J. Cell. Sci. *124,* 4299-4308.

Serbus, L.R., and Sullivan, W. (2007). A cellular basis for *Wolbachia* recruitment to the host germline. PLoS Pathog. *3,* e190.

Shen, R., Weng, C., Yu, J., and Xie, T. (2009). eIF4A controls germline stem cell self-renewal by directly inhibiting BAM function in the *Drosophila* ovary. Proc. Natl. Acad. Sci. U. S. A. *106,* 11623-11628.

Sheng, X.R., Brawley, C.M., and Matunis, E.L. (2009). Dedifferentiating spermatogonia outcompete somatic stem cells for niche occupancy in the *Drosophila* testis. Cell. Stem Cell. *5,* 191-203.

Shibata, M., Blauvelt, K.E., Liem, K.F.,Jr, and Garcia-Garcia, M.J. (2011). TRIM28 is required by the mouse KRAB domain protein ZFP568 to control convergent extension and morphogenesis of extra-embryonic tissues Development *138,* 5333-5343.

Song, X., Wong, M.D., Kawase, E., Xi, R., Ding, B.C., McCarthy, J.J., and Xie, T. (2004). Bmp signals from niche cells directly repress transcription of a differentiation-promoting gene, *bag of marbles,* in germline stem cells in the *Drosophila* ovary. Development *131,* 1353-1364.

Song, X., and Xie, T. (2002). DE-cadherin-mediated cell adhesion is essential for maintaining somatic stem cells in the *Drosophila* ovary. Proc. Natl. Acad. Sci. U. S. A. *99,* 14813-14818.

Song, X., Zhu, C.H., Doan, C., and Xie, T. (2002). Germline stem cells anchored by adherens junctions in the *Drosophila* ovary niches. Science *296,* 1855-1857.

Spradling, A., Fuller, M.T., Braun, R.E., and Yoshida, S. (2011). Germline stem cells Cold Spring Harb Perspect. Biol. *3,* a002642.

Starr, D.J., and Cline, T.W. (2002). A host parasite interaction rescues *Drosophila* oogenesis defects. Nature *418,* 76-79.

Storz, J.F., and Wheat, C.W. (2010). Integrating evolutionary and functional approaches to infer adaptation at specific loci Evolution *64,* 2489-2509.

Storz, J.F., and Zera, A.J. (2011). Experimental approaches to evaluate the contributions of candidate protein-coding mutations to phenotypic evolution Methods Mol. Biol. *772,* 377-396.

Sun, S., and Cline, T.W. (2009). Effects of *Wolbachia* infection and ovarian tumor mutations on *Sex-lethal* germline functioning in Drosophila. Genetics *181,* 1291-1301.

Swanson, W.J., Clark, A.G., Waldrip-Dail, H.M., Wolfner, M.F., and Aquadro, C.F. (2001). Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. *98,* 7375-7379.

Swanson, W.J., and Vacquier, V.D. (2002). The rapid evolution of reproductive proteins Nat. Rev. Genet. *3,* 137-144.

Swanson, W.J., and Vacquier, V.D. (1995). Extraordinary divergence and positive Darwinian selection in a fusagenic protein coating the acrosomal process of abalone spermatozoa Proc. Natl. Acad. Sci. U. S. A. *92,* 4957-4961.

Swanson, W.J., Wong, A., Wolfner, M.F., and Aquadro, C.F. (2004). Evolutionary expressed sequence tag analysis of Drosophila female reproductive tracts identifies genes subjected to positive selection Genetics *168,* 1457-1465.

Swanson, W.J., Yang, Z., Wolfner, M.F., and Aquadro, C.F. (2001). Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals Proc. Natl. Acad. Sci. U. S. A. *98,* 2509-2514.

Szakmary, A., Cox, D.N., Wang, Z., and Lin, H. (2005). Regulatory relationship among *piwi*, *pumilio*, and *bag-of-marbles* in Drosophila germline stem cell self-renewal and differentiation. Curr. Biol. *15,* 171-178.

Szakmary, A., Reedy, M., Qi, H., and Lin, H. (2009). The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in Drosophila melanogaster J. Cell Biol. *185,* 613-627.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism Genetics *123,* 585-595.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0 Mol. Biol. Evol. *24,* 1596-1599.

Tazuke, S.I., Schulz, C., Gilboa, L., Fogarty, M., Mahowald, A.P., Guichet, A., Ephrussi, A., Wood, C.G., Lehmann, R., and Fuller, M.T. (2002). A germline-specific gap junction protein required for survival of differentiating early germ cells Development *129,* 2529-2539.

Teixeira, L., Ferreira, A., and Ashburner, M. (2008). The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. PLoS Biol. *6,* e2.

Thornton, K., and Andolfatto, P. (2006). Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of *Drosophila melanogaster*. Genetics *172,* 1607-1619.

Tsaur, S.C., Ting, C.T., and Wu, C.I. (1998). Positive selection driving the evolution of a gene of male reproduction, *Acp26Aa,* of Drosophila: II. Divergence versus polymorphism. Mol. Biol. Evol. *15,* 1040-1046.

Turner, L.M., and Hoekstra, H.E. (2008). Causes and consequences of the evolution of reproductive proteins Int. J. Dev. Biol. *52,* 769-780.

Veneti, Z., Bentley, J.K., Koana, T., Braig, H.R., and Hurst, G.D. (2005). A functional dosage compensation complex required for male killing in Drosophila Science *307,* 1461-1463.

Venken, K.J., He, Y., Hoskins, R.A., and Bellen, H.J. (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. Science *314,* 1747-1751.

Wang, Z., and Lin, H. (2004). Nanos maintains germline stem cell self-renewal by preventing differentiation Science *303,* 2016-2019.

Watts, T., Haselkorn, T.S., Moran, N.A., and Markow, T.A. (2009). Variable incidence of *Spiroplasma* infections in natural populations of Drosophila species. PLoS One *4,* e5703.

Werren, J.H., Baldo, L., and Clark, M.E. (2008). *Wolbachia*: master manipulators of invertebrate biology. Nat. Rev. Microbiol. *6,* 741-751.

Wong, A., Turchin, M., Wolfner, M.F., and Aquadro, C.F. (2012). Temporally variable selection on proteolysis-related reproductive tract proteins in Drosophila Mol. Biol. Evol. *29,* 229-238.

Wong, A., Turchin, M.C., Wolfner, M.F., and Aquadro, C.F. (2008). Evidence for positive selection on *Drosophila melanogaster* seminal fluid protease homologs. Mol. Biol. Evol. *25,* 497-506.

Wong, M.D., Jin, Z., and Xie, T. (2005). Molecular mechanisms of germline stem cell regulation. Annu. Rev. Genet. *39,* 173-195.

Wu, M., Sun, L.V., Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J.C., McGraw, E.A., Martin, W., Esser, C., Ahmadinejad, N., *et al.* (2004). Phylogenomics of the reproductive parasite *Wolbachia pipientis w*Mel: a streamlined genome overrun by mobile genetic elements. PLoS Biol. *2,* E69.

Xi, Z., Gavotte, L., Xie, Y., and Dobson, S.L. (2008). Genome-wide analysis of the interaction between the endosymbiotic bacterium *Wolbachia* and its Drosophila host. BMC Genomics *9,* 1.

Xia, L., Jia, S., Huang, S., Wang, H., Zhu, Y., Mu, Y., Kan, L., Zheng, W., Wu, D., Li, X., *et al.* (2010). The Fused/Smurf complex controls the fate of Drosophila germline stem cells by generating a gradient BMP response Cell *143,* 978-990.

Xie, T., Song, X., Jin, Z., Pan, L., Weng, C., Chen, S., and Zhang, N. (2008). Interactions between stem cells and their niche in the Drosophila ovary Cold Spring Harb. Symp. Quant. Biol. *73,* 39-47.

Xie, T., and Spradling, A.C. (2000). A niche maintaining germ line stem cells in the Drosophila ovary Science *290,* 328-330.

Xie, T., and Spradling, A.C. (1998). *decapentaplegic* is essential for the maintenance and division of germline stem cells in the Drosophila ovary. Cell *94,* 251-260.

Yamada, R., Floate, K.D., Riegler, M., and O'Neill, S.L. (2007). Male development time influences the strength of *Wolbachia*-induced cytoplasmic incompatibility expression in *Drosophila melanogaster*. Genetics *177,* 801-808.

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood Mol. Biol. Evol. *24,* 1586-1591.

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood Comput. Appl. Biosci. *13,* 555-556.

Yue, L., and Spradling, A.C. (1992). *hu-li tai shao*, a gene required for ring canal formation during Drosophila oogenesis, encodes a homolog of adducin. Genes Dev. *6,* 2443-2454.