

Spring 5-1-2011

# Identification, cloning and fluorescent-tagging of protamines of the flour beetle, *Tribolium castaneum*

Kristina Martimucci

Follow this and additional works at: [https://surface.syr.edu/honors\\_capstone](https://surface.syr.edu/honors_capstone)



Part of the [Biology Commons](#), and the [Ecology and Evolutionary Biology Commons](#)

---

## Recommended Citation

Martimucci, Kristina, "Identification, cloning and fluorescent-tagging of protamines of the flour beetle, *Tribolium castaneum*" (2011).  
*Syracuse University Honors Program Capstone Projects*. 266.  
[https://surface.syr.edu/honors\\_capstone/266](https://surface.syr.edu/honors_capstone/266)

This Honors Capstone Project is brought to you for free and open access by the Syracuse University Honors Program Capstone Projects at SURFACE. It has been accepted for inclusion in Syracuse University Honors Program Capstone Projects by an authorized administrator of SURFACE. For more information, please contact [surface@syr.edu](mailto:surface@syr.edu).

**Identification, cloning and fluorescent-tagging of protamines of  
the flour beetle, *Tribolium castaneum***

A Capstone Project Submitted in Partial Fulfillment of the  
Requirements of the Renée Crown University Honors Program at  
Syracuse University

By

Kristina Martimucci

Candidate for B.S. Degree

And Renée Crown University Honors

May 2011

Honors Capstone Project in Biology

Capstone Project Advisor: \_\_\_\_\_  
Dr. John Belote

Honors Reader: \_\_\_\_\_  
Dr. Eleanor Maine

Honors Director: \_\_\_\_\_  
Dr. James Spencer,  
Interim Director

Date: \_\_\_\_\_

## ABSTRACT

The genome of eukaryotes is associated with chromosomal proteins that fold the DNA into ordered arrangements. The most abundant of these are the histones. Four of these (H2A, H2B, H3 and H4) form the nucleosomal core around which DNA is wrapped. Other proteins can associate with chromatin and fold it into higher order structures. During spermatogenesis in many organisms the standard chromatin configuration undergoes a transition to a highly condensed form in which small, highly basic proteins called protamines replace the histones. It is thought that protamine-based chromatin is more resistant to mutational damage.

Previous work by the Renkawitz-Pohl group identified and cloned the protamine genes from *Drosophila melanogaster*. In this species there are two tandemly arranged genes encoding similar small, highly basic proteins, ProtamineA and ProtamineB, that are expressed in late spermatogenesis and that accumulate in the sperm head as the histones disappear. Interestingly, these protamines are not essential for sperm function, although sperm lacking the protamines were more susceptible to mutation.

In the present study I sought to extend these studies to another insect species, the red flour beetle *Tribolium castaneum*. Here I describe the identification of three clustered genes (called Protamine-1, Protamine-2, and Protamine-3) encoding small highly basic proteins that show similarity to the *Drosophila* protamines. To see if these are indeed protamines, I have used gene splicing methods to tag each of them with a fluorescent marker (GFP, a

green fluorescent protein) and have subcloned the tagged constructs into a germline transformation vector, *pB3xP3-EGFP*, that will be used to create transgenic beetles that express the fluorescent proteins. To date, Dr. Belote has generated transformed lines of fluorescently tagged Protamine-1 and has confirmed that this gene is expressed during spermatogenesis and that the protein accumulates in the sperm head as expected for a protamine. I have successfully GFP-tagged the Protamine-2 gene and subcloned it into the transformation vector. The next step will be to generate transgenic beetles and examine the sperm-specific expression of this candidate protamine gene. Generation of the Protamine-3 GFP-tagged construct is in progress.

It is anticipated that these fluorescent-tagged protamine lines will be an invaluable genetic tool for a variety of studies in which direct visualization of live sperm is desired.

*This work was supported by a Ruth Meyer Award for Undergraduate Research  
and a Renee Crown Honors Program Research Award*

## TABLE OF CONTENTS

INTRODUCTION.....	5
MATERIALS AND METHODS.....	25
RESULTS.....	35
DISCUSSION AND FUTURE DIRECTION.....	68
REFERENCES.....	71
WRITTEN CAPSTONE SUMMARY.....	75

## INTRODUCTION

### **Chromatin and chromosomal proteins**

The genome of eukaryotic cells does not exist as extended “naked” DNA molecules, but rather is highly folded and condensed, and is associated with chromosomal proteins organized into a nucleoprotein assembly called chromatin (Lewis, et al., 2003). This nucleoprotein assembly, consisting of DNA interacting with histone and non-histone proteins, confers several levels of ordered structure. The basic subunit of chromatin is the nucleosome, which consists of 146 (bp) of DNA wrapped around a histone (H3-H4) tetramer associated with 2 (H2A-H2B) histone heterodimers. Another histone called HI, or linker histone, associates with the DNA that lies between adjacent nucleosomes. These particles are dynamically active in maintaining higher levels of chromatin organization. That is, the nucleosomal “beads on a string” configuration can be further folded into a 30 nm chromatin fiber, as well as more complex looped domains along the chromosomes (Fig. 1).

As “DNA packaging” proteins, histones have been historically considered to play a purely structural role in the mediation of chromatin structure. However, more recent studies have revealed that exchange of specialized histone variants and the modification of core histone tails, which extend outward from the nucleosome perimeter, provide complex levels of epigenetic regulation affecting cellular processes such as transcription and DNA repair (Henikoff, 2008). Thus,

Fig 1.

histone variability and modification can alter the structural nature of nucleosomes to help in the control of gene function.

### **Protamines: Sperm-specific chromosomal proteins**

In many species, the chromatin structure typical of somatic cells undergoes a dramatic compaction during sperm development (see Lewis, et al., 2003 for review). This transition is correlated with the appearance of sperm-specific chromosomal proteins called protamines. It is envisioned that during the late stages of spermatogenesis, the histones that make up nucleosomal-based chromatin are removed, and a protamine-based, highly condensed, chromatin structure is assembled. This switch is thought to be facilitated by additional chromosomal proteins, such as the “protamine-like” proteins, and various histone “variant” proteins, as well as others (Rathke, et al., 2007).

Unlike histones, protamines from different organisms are not highly conserved in their primary amino acid (aa) sequences; however they do have similar overall properties (Balhorn, 2007). That is, they are characterized as highly basic nuclear proteins that are typically quite short (50-110 aa in length). For example, protamine-2 of humans is a 102 aa protein with a preponderance of basic residues, especially in the C-terminal half of the protein (Fig. 2). Two identifying structural elements within the protamine polypeptide have been found: small ‘anchoring’ domains containing multiple positively charged arginine or lysine amino acids (used to bind the protein to DNA); and multiple serine and threonine residues (used as phosphorylation sites) (Balhorn, 2007).



Fig.2

The P1 protamine of placental mammals has also been well studied. It is about 50 aa long, and contains three domains: a central arginine-rich DNA binding domain flanked by short peptide segments containing cysteine residues (Balhorn, 2007). These cysteines could conceivably be involved in disulfide bond formation. This central DNA binding domain typically consists of a series of anchoring sequences containing 3-11 consecutive arginine residues, which bind the protein to the DNA. When comparing this domain between fish protamines and mammalian P1 protamines, there is a high conservation in this anchoring domain, but other than that, there is high variation between the other domains (Balhorn, 2007).

### **Spermatogenesis and histone replacement**

Spermatogenesis is a complex developmental process that involves amplification of germline stem cells, their differentiation into spermatocytes, two meiotic divisions, and finally dramatic differentiation into mature spermatozoa (Fuller, 1998; Olivia and Castillo, 2011). During this process, there is a dramatic nuclear transformation, including chromatin reorganization (Rathke et al., 2010), which is essential for spermatogenesis. A complex cascade of transcriptional and regulatory events facilitates this dramatic reorganization of local and global chromatin structure (Jayaramaiah et al, 2005).

During these late stages of spermatogenesis, where spermatid differentiation occurs, the initially round spermatids undergo extensive morphological changes that give rise to elongated spermatids (Fuller, 1998).

These extensive morphological changes of the sperm can be attributed to two processes: generation of a long sperm tail containing a structure called the axoneme that acts as the motility apparatus, and the elongation of the sperm head, that is associated with significant condensation and elongation of the sperm nucleus. This latter process is thought to be caused, at least in part, by the switch from histone-based chromatin to protamine-based chromatin (Fig. 3). It is hypothesized that protamine binding to DNA in these late stages of spermatogenesis results in the production of an uncharged chromatin complex that enables the DNA molecules to be condensed into a volume some  $1/20^{\text{th}}$  that of a somatic nucleus (Balhorn, 2007). It has been speculated this protamine-based chromatin organization, that leads to a compact hydrodynamic sperm head, can work to protect the paternal genome from physical and chemical damage (Rathke et al., 2010; Jayaramaiah et al., 2005).

In certain plant and animal species, the histones that package DNA in early spermatids are removed from the DNA and replaced in the final stages of spermatid maturation by one of three types of proteins: sperm-specific histones, protamine-like proteins or protamines. Studies have shown that there are three general mechanisms for histone replacement: (a) large-scale incorporation of histone variants, creating less stable nucleosomes; (b) genome-wide histone hyperacetylation; and (c) competition for DNA binding with very basic DNA-interacting non-histone proteins such as transition proteins and protamines (Gaucher, et al. 2010).

Fig. 3

### **Rapid divergence and evolution of protamines**

Comparing various organisms, there is a wide variety in the number and types of genes that code for protamine or protamine-like proteins. For example, in salmonid fish there are 15 closely related protamine genes, while birds have two copies of almost identical protamine genes, and mammals have two very distinct protamine genes, called *protamine-1* and *protamine-2* (Balhorn, 2007). Focusing on the mammalian protamines, it is widely thought that the *protamine-2* gene arose by tandem duplication of *protamine-1*. Both genes are located in a tight cluster on chromosome 16, and they show significant amino acid similarity, although they differ in overall length (*protamine-1* is 50 aa long, and *protamine-2* is 102 aa). It has been shown that *protamine-2* undergoes proteolytic cleavage which removes the N-terminal half, making the processed *protamine-2* similar in size to *protamine-1*.

Interspecific comparisons of protamines have shown that these proteins are very rapidly evolving (Balhorn, 2007). For example, Retief, et al., (1993) compared the *protamine-1* genes from several primate species, including human, and found a remarkably high amino acid substitution rate, among the fastest known for any protein. This rapid evolution is also seen when comparing protamine genes from different species of *Drosophila*. For example, Fig. 4 shows the amino acid alignments of the known protamines from five closely related species of the melanogaster subgroup (*D. melanogaster*, *D. simulans*, *D. sechellia*, *D. yakuba*, and *D. erecta*). While these proteins clearly represent products of orthologous genes, there are significant differences in the sequences. The degree

Fig. 4

of divergence is very high, considering that these species shared a common ancestor less than 10 million years ago (Fig. 5). An even more dramatic evolutionary divergence in protamine sequence is seen when comparing *D.melanogaster* and *D. virilis* (divergence time 40 Mya). As shown in Fig. 6, here the amino acid sequences show a very weak alignment (Fig. 6). This rapid evolutionary divergence makes it difficult to identify orthologous protamine genes by database searches when comparing species that are not close relatives.

Protamine-like proteins have also been found in many species, and are characterized by a high lysine and arginine content (35-50% Arg and Lys), comparable to that seen in histones, and are considerably larger than true protamines (up to 300 aa's). These structural characteristics showing similarity to both histones and protamines suggest that protamine-like proteins represent an evolutionary mid-transition to modern protamines (Balhorn, 2007). Evolution to a true protamine is hypothesized to occur as the protamine-like gene becomes better able to influence the stability of the sperm chromatin complex, and exhibit more efficient success in displacement of histones and transition proteins from DNA, as a true protamine gene will accomplish.

### **Studies of *Drosophila* protamines**

Although many studies have been done in mammalian systems to investigate the mechanisms of the histone to protamine transition, less is known for the *Drosophila melanogaster* model. To date, much of the information about *Drosophila* protamines has been obtained by the continued studies of the

Fig.5



Fig. 6

Renkawitz-Pohl group (Jayaramaiah-Raja and Renkawitz-Pohl, 2005; Rathke, et al., 2007; Rathke, et al., 2010). Jayaramaiah-Raja and Renkawitz-Pohl (2005) were the first to identify the protamine genes in this species, and show that the histone-protamine transition occurs during the late stages of spermatogenesis. Analysis of the *Drosophila* genome sequence identified two tandemly arranged genes, called *Mst35Ba* and *Mst35Bb*, that encoded proteins showing similarity to mammalian protamines (Fig. 7). These genes were originally defined in a cDNA library screen as encoding male-specific transcript of unknown function. Because their sequences showed similarity to each other and to known protamine genes from other species (i.e., small, highly basic proteins), it was hypothesized that they represented protamine genes. To test this, the Renkawitz-Pohl group created transgenic flies with GFP-tagged *Mst35Ba* and *Mst35Bb*, and demonstrated that these proteins are expressed in the male germline and replace histones during the late stages of spermatogenesis, as expected for protamines. These genes were therefore renamed *ProtamineA* and *ProtamineB*.

Another specifically male transcribed gene, *Mst77F*, a distant relative of the histone H1/H5 (linker histone) family, was proposed to play a role either as a transition protein or a replacement protein for the compaction of the *Drosophila* sperm chromatin.

While it was known that *Drosophila* histone-based nucleosomal chromatin configurations are replaced by a protamine-based chromatin structure during sperm maturation, the biological advantage of this chromatin reconfiguration is

Fig. 7

unclear. To study the biological effects of the two protamine genes, gene knockouts of the *Mst35Ba* and *Mst35Bb* plus three unrelated (and non-essential) genes were created by the deletion of *Df(2L)FDD-0338164*, referred to as *protΔ* (Rathke, et al., 2010). The observation that flies homozygous for this deletion were fertile, although somewhat reduced, proves that these protamines are not essential for either viability or fertility (although about 20% of the sperm showed abnormal sperm head morphology). This was surprising because previous research had shown that in mice and humans, loss of even one of the two copies of the protamine-1 or protamine-2 genes results in complete male sterility (that is, the protamine genes are haplosufficient). These results showed that in *Drosophila*, the protamines are not absolutely required for sperm differentiation and function. This raised the question of what is the function of protamines. One possibility is that they act to protect the paternal genome from mutagens. This was tested using the Muller5 test (Muller and Altenburg, 1919). In this procedure flies are subjected to X-irradiation, and the frequency of induced recessive lethal mutations provides a measure of mutagen sensitivity. Results showed that Protamine-A and B-deficient sperm are more sensitive to X-rays than wild-type sperm (Rathke, et al., 2010).

### **The use of fluorescently-tagged protamines to study sperm mobility, behavior, and function in *Drosophila***

In addition to being of interest in its own right, the identification and characterization of *Drosophila* protamines has led to the development of a

molecular genetic tool that has provided a breakthrough in the ability to visualize live sperm within the female reproductive tract. Using gene splicing methods, the protamines have been tagged with green or red fluorescent markers (either Green Fluorescent Protein, GFP, or Red Fluorescent Protein, RFP), and transgenic *Drosophila* strains generated in which males produce sperm with green or red sperm heads. Such sperm can be easily identified and tracked inside the dissected female sperm storage organs (Fig. 8). Using these lines, Manier, et al. (2010) were able to address many previously unanswered questions regarding the nature of sperm competition when females mate consecutively with two males, and how sperm move and interact with one another inside the female. For example, sperm were found to exhibit a much higher degree of mobility within the female's sperm storage organs than originally thought, and the sperm remained motile for several days, even after mating with a competitor male. It was also seen that when a female mates with a second male, the stored sperm from the first male have a tendency to leave storage, and then be replaced by sperm from the second male. This observation helps explain the long-known fact that in multiply mated female, most progeny are sired by the last male to mate (a phenomenon known as second male sperm precedence).

A related phenomenon is conspecific male precedence. In this case it has been seen that when a female mates with two males, one of her own species (the conspecific male) and one from a closely related species (the heterospecific male), the female produces a large majority of offspring sired by the conspecific male, regardless of whether he was the first or second to mate. To examine the basis for

Fig. 8

this, the GFP- and RFP- tagged *D. melanogaster* protamine transgenes have been introduced into the sibling species *D. simulans* and *D. mauritiana* so that their sperm can be distinguished in interspecific crosses, and the details of sperm storage and motility inside the dual mated female are now being studied (Manier, et al., in preparation).

Other experiments have involved introducing the *D. melanogaster* tagged protamine genes into more distantly related Drosophila species (*D. pseudoobscura* and *D. bifurca*). In these cases, the expression of the heterologous *melanogaster* gene is not very strong and the sperm heads, while visible, are only weakly labeled (J. Belote, unpublished results). This could be related to the fact that protamines are rapidly evolving and so the *melanogaster* gene (or its protein gene product) is not expressed well in these other species. Thus, it seems that if this overall approach to visualizing living sperm is to be carried out with other organisms, it might be necessary to tag the protamine gene naturally found in that species or a close relative.

**This research project: Identify, clone and fluorescently tag protamine genes from the flour beetle, *Tribolium castaneum***

While there have been numerous studies of protamines in fish, birds, and mammals, relatively little is known about protamines and the histone-protamine transition in insects. The only detailed studies in insects are those described above for *Drosophila*. To obtain a more complete picture of the structure and function of insect protamines, it would be useful to identify and analyze the protamine genes

of another insect species. Among the types of questions that could be then addressed are: (1) How many protamine genes are there? (2) What is the level of amino acid identity? (3) Are protamines dispensable for fertility in insects other than *Drosophila*? (4) Is the timing of the histone-protamine transition during spermatogenesis well conserved in insects? (5) Are there recognizable regulatory sequences that specify testis-specific expression of protamines?

As a starting point to get at these questions, the present study was undertaken to identify, clone, and GFP-tag the protamine genes from the red flour beetle, *Tribolium castaneum*. This species was chosen because of the many genetic tools that are available to manipulate its genome (Denell, 2008). For example, its genome has been fully sequenced and is available in a public database, its genetics and development are well studied, there are many mutant strains available from a stock center, it is easy to rear in the lab, and an effective germline transformation protocol has been worked out. In addition, *Tribolium* and *Drosophila* are far enough apart on the evolutionary tree (they diverged about 280 Mya) that conserved DNA and amino acid sequences should be meaningful.

A separate goal of this work was to create *Tribolium* strains that express green or red fluorescent sperm heads so that studies of sperm motility, behavior and competition, similar to those done with *Drosophila*, could be extended to the beetle model system. There have been numerous studies of post-copulatory sexual selection in *Tribolium* (Bloch Qazi, 2003; Michalczyk, et al., 2010), but the mechanisms by which sperm from different males are favored, and what goes on inside the female's reproductive tract are unknown. The material produced in this



present study (e.g., transgenic beetles with green and/or red labeled sperm heads) should prove invaluable for investigating these topics in the beetle model system.

## MATERIALS AND METHODS

### ***Tribolium castaneum* culture maintenance**

Culturing and handling of *Tribolium* were done according to the procedures described in Berghammer, et al. (1999) and Sokoloff (1974), as well as those described on the “Beetle Wrangling” website of Prof. Richard Beeman (USDA Agricultural Research Service, Manhattan, KS):

<http://bru.gmpcr.ksu.edu/prof/tribolium/wrangle.asp>. All *Tribolium castaneum* lines (the wild-type GA-2 and the mutant *pearl* stocks) were obtained from the Tribolium Stock Center (USDA-ARS-GMPRC, Manhattan, KS) and maintained in a medium of organic whole grain wheat flour (Bob’s Red Mill, Milwaukie, Oregon), enriched with 5% torula yeast powder (USB Corp., Cleveland, OH), and (as anti-fungal agent) 0.3 g of Fumagilin-B (Medivet Pharmaceuticals, Alberta, Canada) per kilogram of flour. Cultures were kept in 6 oz. polypropylene *Drosophila* bottles or in polystyrene vials (25 mm x 95 mm) in a humid incubator at 31° C.

### **Isolation of DNA from Adult *Tribolium***

Approximately 50 beetles (Strain GA-2) were added to a 15mL Dounce homogenizer along with 5 mL Homogenization Buffer (HB, 10mM Tris HCl (pH 7.5), 60 mM NaCl, 10mM EDTA, 150µM spermine, 150µM spermidine). The beetles were ground with the A pestle (looser fit than the B pestle in order to not rupture the nuclei). The mixture was transferred to a 14mL Sarstedt centrifuge

tube and spun in a Sorvall RC5 centrifuge at 1000rpm for one minute to pellet the cuticle and beetle debris but not the nuclei. The supernatant was removed with a transfer pipette and transferred to a new 14mL Starstedt tube. This was centrifuged at 8000 rpm for 5 minutes to pellet the nuclei.

The brown pellet was resuspended in 0.5mL HB and transferred to a new 1.5 mL Eppendorf microfuge tube. Proteinase K was added to a final concentration of 100  $\mu\text{g}/\text{mL}$  and mixed. 50 $\mu\text{L}$  of 10% SDS was added and mixed by inverting several times, and the tube incubated for 45-60 minutes at 37°C.

Next, 0.5 mL of phenol/chloroform/isoamyl alcohol (25:24:1) was added and the mixture was vortexed vigorously, spun for a few minutes, and then the aqueous top layer was transferred to a new 1.5 mL tube. This phenol/chloroform/IAA extraction was repeated, and the aqueous layer was transferred to a new tube. A small amount of H<sub>2</sub>O was added to bring the volume back to 0.5mL and then 0.5mL of chloroform was added, mixed, and spun. The aqueous layer (top) was then transferred to a new 1.5 mL tube.

20 $\mu\text{L}$  of 5M NaCl was then added to the tube and after mixing, 1.0 mL of absolute ethanol was slowly added to the tube, by layering it on top. The tube was gently mixed by inverting several times. At this point, the DNA precipitated and appeared as a fluffy, fibrous, tangle of strands. The precipitate was brownish because of the pigments that co-purify with DNA.

A glass hook was used to transfer the DNA to a new 1.5 tube containing 0.5 mL of 70% ethanol. The tube containing the DNA and the 70% ethanol was then spun for 5 minutes to pellet the DNA. The ethanol was then poured off and

the sides were blotted dry with a Kimwipe. The tube was vacuum-dried for 10 min and the DNA was resuspended in 100  $\mu$ L of TE containing RNaseA (20 $\mu$ g/mL). This sample was incubated at 37°C for 20 minutes with occasional light vortexing to dissolve the DNA.

Next, 260  $\mu$ L of TE and 40 $\mu$ L 3M NaOAc was added, and mixed. 1.0 mL of absolute ethanol was added, mixed, and the tube was placed at -20°C for 20 minutes. After microfuging at 13000 rpm for 10 minutes, the ethanol was poured off, and the pellet washed with 70% ethanol. The DNA pellet was vacuum-dried for 10-20 min, resuspended in 100  $\mu$ L of TE at 37°C with occasional light vortexing, and left overnight in the refrigerator (about 4°C) to allow for the pellet to completely dissolve.

The quality of the DNA was checked by running 3 $\mu$ L of uncut DNA, and 3  $\mu$ L of *Eco*RI cut DNA, on an agarose gel. For uncut DNA, a prominent high molecular weight band is expected, with little smearing (which indicated DNA degradation). The restriction digest should give a broad smear, indicating that the DNA is not resistant to enzymatic digestion.

### **Synthesis of Oligonucleotide Primers**

DNA primers required for PCR and Site-directed Mutagenesis were ordered from Sigma-Genosys (<http://www.sigmaaldrich.com/life-science/custom-oligos.html>). All samples were received in a dried state and were made into a 250 $\mu$ M stock solution by dissolving in TE. For working solutions of primers, the stock solution was typically diluted to 5  $\mu$ M with TE.

The following primers were used to PCR amplify the fragments of the *Tribolium* genome containing the protamine-like genes:

Primer #	Sequence
Primer 2	CAGTTAGCTTCGGTCCGAAATG
Primer 4	AGCACATCAAAATCTATAAGATAG
Primer 5	TCTCCAGTTACGCATCTTCTAACCTCGT
Primer 8	GGGCCCTCGTCAAAATCATTCATGC
Primer 7	TCCCGAAGCTACGCTACAGCTATTAA
Primer 9	TAATGGGCAAAGGAGTTTTCTGT

### **Polymerase Chain Reaction**

The PCR reaction typically contained 4 $\mu$ l 5x GoTaq Polymerase Reaction Buffer (Promega, Madison, WI), 2 $\mu$ l 200  $\mu$ M dNTP mixture (containing equimolar amounts of dATP, dCTP, dGTP, dTTP), 1 $\mu$ l 5' Primer (5  $\mu$ M), 1 $\mu$ l 3' Primer (5 $\mu$ M), 1 ng of *Tribolium* genome DNA, and dH<sub>2</sub>O brought up to 20  $\mu$ l. The initial denaturing step was performed at 94° for 5 minutes, 5 $\mu$ l of diluted Taq polymerase was added (1 part Taq polymerase, 10 parts 5X PCR buffer, 40 parts water) and the sample was amplified for 35 cycles (94°C for 1 minute, 58°C for 1 minute, 72°C for 1 minute) and completed with a final extension period at 72°C for 8 minutes. The samples (25  $\mu$ L total volume) were then stored at 4°C; a 15  $\mu$ L aliquot was checked by agarose gel electrophoresis.

### **TOPO TA Cloning of PCR Products**

The TOPO TA Cloning kit by Invitrogen was used to clone the PCR product into a suitable vector using Topoisomerase I. The reaction was as follows: 0.5 to 4 $\mu$ l of fresh PCR product, 1 $\mu$ l of 1/4 diluted salt solution (1X=1.2 m NaCl<sub>2</sub>, .06 m MgCl<sub>2</sub>), water to bring the reaction up to 5 $\mu$ l, 1 $\mu$ l of TOPO vector. The reaction was mixed gently and incubated for 20 min at room temperature (22-

23°C). The reaction was placed on ice, and the *E. coli* transformed by electroporation or heat shock-method.

### **Transformation of *E. coli***

Two different methods were used for transformation: electroporation or heat-shock method. For the transformation of the first PCR product containing the *Tribolium* genes (with primers 8/5 and 7/9), into the TOPO vector, electroporation was used. Electrocompetent cells were made using *E. coli* strain DH5 $\alpha$ . The cells were collected during the exponential growth phase, then washed repeatedly with 10% glycerol to remove soluble compounds (e.g., salts) detrimental to the electroporation. 80 $\mu$ l aliquots of these cells were quick frozen in a dry ice/ethanol bath and stored at -70° C until needed. For transformation, cells were thawed on ice and 2 $\mu$ L of ligation mixture (2 $\mu$ L DNA, 1 $\mu$ L dil salt, 1 $\mu$ L TOPO vector, 2 $\mu$ L sterile H<sub>2</sub>O) was added to the cell aliquots. Cells were placed in a cold electroporation cuvette and subjected to 1800 V in an Eppendorf 2510 Electroporator. Cells were immediately transferred to 800 $\mu$ l SOC media in a 14 ml sterile culture tube and incubated with shaking for 45 minutes at 37°C. 300 $\mu$ l of the mixture was then placed in a 5 mL sterile tube. At this point, the electroporation transformation used the same procedure as the standard transformation from the addition of antibiotic onward.

The standard heat shock transformation of *E. coli* used DH5 $\alpha$  chemically-competent *E. coli* cells which were stored at -70°C, and thawed on ice before use. 5 $\mu$ L of the ligation mixture or a 100-fold dilution of purified plasmid DNA was added to the cells, which were incubated on ice for 20 minutes. The mixture of

cells and DNA was then given a 90 second heat shock in a 42°C heating block and returned to the ice. The mixture was transferred to a 5 mL sterile tube and 200 µL of LB was added. The tubes were incubated 1-2 hours at 37°C. At the end of the incubation, 20µL of 20 mg/mL ampicillin was added to the mixture. 50µL of 20% X-gal and 10µL 100mM IPTG (isopropyl-thio-galactosidase) was added to the mixture as well to allow blue/white selection of colonies of interest (i.e., colonies containing plasmid only turn blue, colonies containing recombinant plasmids are white). 3mL of top agarose was added to the 37°C tube, and the top agar spread evenly over pre-warmed LB-Amp plates. After 10 minutes, the plates were inverted and placed in a 37°C incubator overnight.

### **Isolation of Plasmid DNA from *E.coli***

Two protocols were used to isolate plasmids from *E.coli*: Plasmid Miniprep method and Wizard Plasmid Prep method. The former was used for most general purposes while the latter was used when high-purity DNA was needed (e.g., for DNA sequencing).

#### **(1) Plasmid Miniprep Method**

In a 14mL sterile tube, 1.5 mL of TB (Terrific Broth) containing Ampicillin (50 µg/mL) was inoculated with a colony of transformed *E.coli* cells that were selected on LB+Amp plates. The tubes were allowed to shake overnight in a 37°C incubator.

The cultures were then poured into 1.5 mL microfuge tubes and spun for 3 min to pellet the cells. The liquid was poured off, 200 µL of Resuspension Buffer

(Promega) was added, and the pelleted cells resuspended by vortexing. 200  $\mu$ L of Lysis Buffer (Promega) was added, and the sample mixed by inversion about 5 times. After 5 min, 200  $\mu$ L of Neutralizing Buffer (Promega) was added, and the sample mixed by inversion several times. The tubes were put on ice for 10 minutes, and then spun in the microfuge at 13000 rpm for 10 minutes. The supernatant was transferred to new 1.5 mL tubes using a pipette, using care not to transfer the white precipitate. 450  $\mu$ L of isopropanol was added to the supernatant, the tubes were allowed to sit on ice for 10 minutes and then spun for another 10 minutes. The isopropanol was poured off, and 500 $\mu$ L of cold 70% ethanol was added to wash the DNA pellets. The ethanol was removed, and the tubes were vacuum dried for about 10 minutes. The DNA pellets were resuspended in 100  $\mu$ L of TE, and incubated at 37°C for 10-15 minutes with occasional vortexing to allow the DNA to dissolve. The amount and integrity of the DNA was then checked by restriction digestion and gel electrophoresis.

## **(2) Wizard Plasmid Prep**

The Wizard Plasmid Purification Kit (Promega, Madison, WI) was used for high yields of quality DNA. The procedure was identical to that described above, up to the point where the Neutralizing Solution was added and the white precipitate spun down. In this procedure, instead of isopropanol precipitation, the supernatant was passed over a Wizard Plasmid Prep purification column.

The cartridges and syringe barrels were set up on the vacuum manifold hooked up to the house vacuum line. 1.0 mL of Wizard Resin (shaken well) was added to each syringe barrel. About 700  $\mu$ l of the supernatant was transferred to



the syringe barrel, where the plasmid DNA bound to the Wizard Resin. The vacuum was then turned on, and the liquid was down through the cartridge, while the resin with the bound DNA was retained. The stopcocks were then closed and 2.0 mL of Wash Solution from the Wizard Kit was added.

For each of the samples, the cartridge with the DNA and the resin were placed on top of a 1.5 mL tube with its top removed. These were placed in the microfuge and briefly spun at 13000 rpm to dry the resin. Another set of 1.5 Eppy tubes were prepared by cutting the tops off. The now dry cartridges were placed on top of the new tubes and 50  $\mu$ L of 50° C TE was added. After sitting for a couple of minutes to elute the DNA the tubes were centrifuged for 5 min, and the DNA solution transferred to a 500  $\mu$ L sterile tube. The DNA was checked by gel electrophoresis, with the average yield being approximately 300-500  $\mu$ g/mL.

### **Restriction Enzyme Digestion and Agarose Gel Electrophoresis**

Restriction enzyme digests were normally performed in a 25  $\mu$ L reaction with 10-13 $\mu$ L of distilled H<sub>2</sub>O, 3-6  $\mu$ L of DNA, 2  $\mu$ L 10xbuffer (this is specific for the enzyme used), 1 $\mu$ L endonuclease. In some cases of digesting a cloning vector, 1 $\mu$ L of Shrimp Alkaline Phosphatase (SAP) was added to prevent subsequent reclosing of the plasmid vector during ligation. Restriction digests were incubated at 37°C for a minimum of 1 hour and a maximum of 18 hours.

Gel electrophoresis was typically performed using 0.7% agarose in 1X TAE buffer [40mM Tris-acetate, 2mM ethylene diamine tetraacetic acid (EDTA), pH 8.0] with 0.5  $\mu$ g/mL ethidium bromide. Bromophenol blue (2 $\mu$ L per well) was

used as loading dye. 1X TAE buffer was poured to cover the gel before the DNA was added to the wells. The gel was run at 60-90 V with times ranging from 45-150 minutes, depending on fragment length and gel density. DNA was visualized with a UV transilluminator, and DNA ladder (DNA Hi-Low Marker) was used for approximate size comparison.

### **Ligation into Plasmid Vectors**

Ligation into plasmid vectors was typically done with 6 $\mu$ L of insert DNA (approximately 100-200 ng), 3 $\mu$ L plasmid vector (approximately 50-100 ng), 1.5  $\mu$ L 10X Ligation Buffer, 1 $\mu$ L T4 DNA ligase (Promega), and 3.5  $\mu$ L distilled water. Ligation reactions were incubated at 4°-14°C overnight, or for at least 15 hours. Ligation reactions with relatively large insert DNA fragments were subject to an extra hour of room temperature incubation following addition of 1  $\mu$ L additional T4 Ligase. A sample of the ligation mixture was transformed into DH5 $\alpha$  *E. coli* cells using the methods described above.

### **DNA Sequencing**

DNA sequencing was outsourced to GeneWiz (<http://www.genewiz.com>). DNA samples were prepared according to the GeneWiz protocol and submitted via the drop box at SUNY-Upstate. Results were provided online and in text format.

## Site Directed Mutagenesis

Oligo Name	Dimer Sequence (5'-3')
Tc5-8 mutS	CCGAAAAAACGCCGCATATGTTAAACACAGTTCCG
Tc5-8 mutAS	CGGAACTGTGTTTAAACATATGCGGCGTTTTTTCGG

Mutagenesis was performed using the QuickChange Site-directed Mutagenesis Kit (Agilent Technologies, Inc., Santa Clara, CA). Typical reaction conditions and procedure were: 5 $\mu$ L 10x reaction buffer (*Pfu* DNA pol buffer), 1 $\mu$ L of plasmid prep DNA (~20ng), 2 $\mu$ L 10mM primer Sense, 2 $\mu$ L 10mM primer Antisense, 1 $\mu$ L dNTP mix, 38 $\mu$ L distilled water and 1 $\mu$ L *Pfu* Turbo (DNA polymerase) were placed in the PCR machine and run under the following conditions: 95°C for 30 seconds followed by 12 cycles of 95°C for 30 seconds, 55°C for 1 minute, and 68°C at 2 minutes per kb of plasmid length (e.g., 12 minutes for 5.6 Kb plasmid). After the mutagenesis step, 2 $\mu$ L *DpnI* was added and incubated for 2 hours to digest parental DNA plasmid. A sample of the reaction mixture was then subjected to electroporation transformation as described above.

## RESULTS

### ***Tribolium castaneum* as a genetic model system**

To better understand this research project, a review of the *Tribolium* genetic model system would be helpful. Male and female adult beetles are a few millimeters long and are very similar in appearance, but can be distinguished by a small patch of bristles, called the sex patch, on the front legs of the male. Pupae are more easily sexed since males and females have very different genital papillae (Fig.9). The beetle will go through three developmental stages before adulthood is reached, the entire cycle from egg to fertile adult taking about 32 days when reared at 30°C. The embryonic stage is about three days; the larval stage is the longest, about 20 days; and the pupal stage lasts 4 days before the beetle reaches adulthood (see Fig. 10).

The method for producing transgenic beetles is very similar to that used for *Drosophila* (Lorenzen, et al., 2003). That is, the DNA of interest is cloned into a piggybac transformation vector (Horn and Wimmer, 2000) which is mixed with “helper” plasmid encoding the piggybac transposase gene, and early embryos are injected in their germline region within a few hours of egg laying. The survivors are then mated and progeny scored for the transformation vector marker (green fluorescent eyes in the case of the p3xP3/EGPFaf vector).

Fig. 9

Fig. 10

## **Overview of the experimental plan to clone and tag the *T. castaneum* protamine genes**

The overall strategy to identify, clone and tag *Tribolium* protamines is illustrated in Fig 11. Briefly, I used bioinformatics to identify the DNA sequences in the *Tribolium* genome sequence database that showed recognizable similarity to the known *Drosophila* protamine coding sequences, and that had properties expected for a protamine (i.e., small size, highly basic charge). PCR was then used to amplify these candidate protamine genes and they were cloned into a plasmid vector, pCR2.1 TOPO. After subcloning into a smaller vector, pBS/2xAsc, site-directed mutagenesis was used to create a unique *NdeI* restriction enzyme cutting site at the C terminus of the coding sequence. An in-frame *NdeI* GFP cassette was inserted to create a protamine-GFP fusion gene. Next, the GFP-tagged construct was cut out with *AscI* and inserted into the unique *AscI* site of the transformation vector piggybac/3xP3-EGFPaf. This final construct can then be used to introduce the protamine –GFP construct into the *Tribolium* genome using the method of piggybac-mediated germline transformation, where transgenic beetles can be identified by the eye-specific GFP marker carried on the vector. If the protamine candidate gene is really a protamine coding gene, then these beetles should produce sperm with glowing green sperm heads.

Fig. 11



### **Bioinformatics: BLAST search for candidate protamine genes in *Tribolium castaneum***

The National Center for Biotechnology Information (NCBI) website (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used to access the BLAST alignment search tool, which searches DNA or protein sequences in the database for meaningful similarities. The sequence of the *Drosophila melanogaster* ProtamineA protein was used to query the translated *Tribolium castaneum* genome sequence using the tblastn search function, to find candidate protamine genes. Three tandemly arranged annotated genes on Linkage Group 4 were found in the *Tribolium* genome that encoded proteins with significant similarity to the *Drosophila* protamine (E values of 2e-06 to 2e-07). These genes were named LOC663849 (or TcGLEAN\_07827), LOC100141730 (or TcGLEAN\_07828), and LOC100141946 (or TcGLEAN\_07670). Fig. 12 shows the sequence alignment between the *Drosophila melanogaster* ProtamineA and these three *Tribolium* proteins. Although the sequence alignment is not very convincing due to the low number of conserved amino acid, the presence of a high proportion of basic amino acids (e.g., lysine (K) and arginine (R)) and the small size of the proteins supports their role as protamine (Fig. 13: A/B map).

### **PCR amplification and cloning of the candidate protamine genes**

The complete genome of *T.castaneum* is available from the BeetleBase database (<http://www.beetlebase.org/>). Therefore, it was possible to design PCR

Fig. 12

Fig. 13

primers flanking each of the three candidate protamine genes so that they could be individually amplified and cloned for analysis. Fig. 14 shows the gene cluster with the positions of the primers for this study. For simplicity, I will refer to Tc07670 as Protamine-1, Tc07827 as Protamine-2 and Tc07828 as Protamine-3. For this experiment genomic DNA was isolated from adult beetles using the protocol described in the Materials and Methods, and then used as a template for PCR.

A series of PCR experiments were carried out, using different combinations of primers to amplify the three genes either individually, or in some cases as a pair. The following combinations of primers were used:

<b>5'-3' Primer</b>	<b>3'-5' Primer</b>	<b>Gene(s) Targeted</b>
Tc9	Tc7	Protamine genes 2&3
Tc9	Tc8	Protamine gene 2
Tc5	Tc8	Protamine gene 2
Tc5	Tc7	Protamine gene 2&3
Tc4	Tc2	Protamine gene 1*

\*Tc 2-4 was done at a separate time therefore not visualized on the same gel

Fig. 15 shows the results of the agarose gel electrophoresis to analyze the products produced by these PCR reactions. The results demonstrated that each of the three genes could be successfully amplified. The sizes of the observed bands fit well with what was expected from the known genomic sequences and the positions of the primers used. For example, primers Tc5 and Tc8 produce the expected 1,260 bp band containing the *Protamine-2* gene region, Tc7 and Tc9 produce the expected 1,200 bp band containing gene *Protamine-2* and *Protamine-3*, and Tc8 and Tc9 produce the expected 810 bp band carrying *Protamine-2*.

Fig. 14

Fig. 15

Although it is not shown on this gel, primers Tc2 and Tc4 yielded a fragment of 1,980 bp that contains *Protamine-1*.

### **Cloning of the PCR products into the pCR2.1 TOPO plasmid vector**

For the next step, I selected the Tc9-7, Tc8-5 and Tc9-8 PCR products to be cloned into the plasmid vector, pCR2.1 TOPO. [Note: Dr. Belote cloned the Tc2-4 fragment into pCR2.1 TOPO.] This linear TOPO vector is very efficient for cloning PCR products because of its ability to ligate DNA fragments containing A-overhangs at their ends using DNA topoisomerase I. The TOPO vector carries a specific sequence, 5'-(C/T)CCTT-3' in addition to a covalently attached topoisomerase I on both of its ends. When the solution of the TOPO vector and PCR product are left to incubate at room temperature with a required diluted salt and Taq Polymerase (which adds an extra adenine nucleotide to the PCR product), the strands will covalently link with high efficiency.

After ligation, the samples were transformed into *E. coli* using electroporation. Colonies successfully grew on the ampicillin plates indicating that they carried the Amp-R gene on the vector. The transformants needed to be checked by isolating plasmid DNA and doing a restriction enzyme digest. The correct banding patterns in each transformant would confirm that Tc9-7, Tc5-8 vector, and the Tc8-9 were successfully inserted into the TOPO vector. In order to do this, six colonies from each plate, a total of 18 colonies, were used to inoculate medium and plasmid DNA was purified from the cultures using a plasmid miniprep procedure. The purified DNA from each of the colonies was then

restriction enzyme digested using the *EcoRI* restriction enzyme. This specific enzyme was used because of the *EcoRI* cutting sites flanking the location of the added PCR product. Because the size of the PCR products for each combination are known, and none of the products contain *EcoRI* sites, the restriction digest should result in two bands. One of the bands should be 3.9 kb, the size of the original TOPO vector without the insert, and the other band should be the size of the original PCR product. The results of this analysis are shown in Fig. 16.

It was concluded that all three PCR cloning experiments were successful. I therefore chose a subset of the confirmed clones to be checked by DNA sequencing, to insure that no PCR-generated mutations had occurred. The samples chosen were:

(TOPO+Tc9-7) # 2 & 3

(TOPO+Tc5-8) # 7 & 8

(TOPO+Tc9-8) # 16 & 17

These DNAs were prepared using the Wizard Prep method, which produces cleaner DNA, and the samples sent to a commercial facility (GeneWiz) for sequencing. Comparison of the sequences with the known sequence in the database revealed that the clones named TOPO/Tc8-5#7 and TOPO/Tc9-7#3 had no mutations and so these were used for subsequent steps.

Although the Tc8-5 and Tc7-9 clones both carried the *Protamine-2* and *Protamine-3* genes, the Tc8-5 fragment had more flanking sequences for *Protamine-2* and Tc7-9 had more flanking sequences for *Protamine-3*. So, Tc8-5 was used for *Protamine-2* analysis and Tc7-9 used for *Protamine-3*. Clones



Fig. 16

carrying the Tc9-8 fragment were not carried forward since they did not provide any additional information that could not be gained from the other clones.

[Note: The further analysis of *Protamine-1*, carried on the cloned Tc2-4 fragment, was done by Dr. Belote. The results of that analysis will be discussed below in the Future Directions section. For the rest of this Results section I will focus on my work with the *Protamine-2* and *Protamine-3* gene clones.]

### **Subcloning of candidate protamine gene fragments into the pBS/2xAsc vector**

While the pCR2.1 TOPO vector is ideal for directly cloning PCR products, it lacks some features that are needed for my subsequent manipulations. Therefore, I wanted to subclone the fragments into another plasmid, pBS/2xAsc, that is more suitable for further work. This vector is smaller than pCR2.1 TOPO (2.9 kb vs. 3.9 kb), which is advantageous for the site-directed mutagenesis step (see below). In addition, this vector has two *AscI* restriction cutting sites flanking the multiple cloning site of pBlueScript so that the subcloned fragments can be cut out with *AscI* for cloning into the piggybac transformation vector's unique *AscI* cloning site (described below).

The pCR2.1 TOPO vector contains *EcoRI* restriction enzyme cutting sites flanking the PCR insert site. So, to remove the PCR insert from TOPO+Tc9-7 and TOPO+Tc5-8, a restriction enzyme digest with *EcoRI* was performed. The

plasmids were also treated with *ScaI*, another restriction enzyme, to cut the TOPO vector sequence into two pieces thereby disabling it from easily re-forming.

Simultaneously, the pBS/2xAsc vector was cut with *EcoRI*, creating a linear plasmid with the ability to be ligated to the insert fragment that had corresponding *EcoRI* cutting site sticky ends. The pBS/2xAsc vector was also treated with Shrimp Alkaline Phosphatase (SAP), which catalyzes the dephosphorylation of the 5' end of the linearized pBS/2xAsc DNA to prevent it from recircularizing without the desired insert.

The next step was to ligate the Tc8-5 and the Tc7-9 DNA fragment inserts with the pBS/2xAsc vector. Two ligation reaction mixes were set up for each. One tube contained T4 ligase, and the other one did not. This was to monitor how well the ligation reaction worked; there should be a considerably smaller number of colonies on the Amp plates for the no ligase reaction, because without ligase the insert fragment and the pBS/2xAsc vector would not be able to join together. After overnight incubation at 14°C, the ligation mixtures were removed, transformed into chemically competent *E. coli* cells, and plated on Amp plates at 37°C. The next day, hundreds of colonies were observed. Colonies from both pBS/2xAsc/Tc5-8 and pBS/2xAsc/Tc7-9 were picked, inoculated into LB media, and after overnight growth at 37°C, plasmid minipreps were done to purify the DNA.

The potential pBS/2xAsc/Tc8-5 clones were worked on first. Several plasmid DNA minipreps were analyzed by *EcoRI* digestion. If the plasmid was the expected pBS/2xAsc/Tc5-8 recombinant, I would have expected to see two bands,

a 2.9 kb band representing the pBS/2xAsc vector, and a 1.3 kb band representing the Tc8-5 fragment. As shown in Fig. 17, plasmid #1 appears to have the correct banding pattern. To confirm that it is the desired clone, I then checked that DNA by cutting it with a number of other restriction enzymes (*ApaI*, *AscI* and *HinCII*). Fig. 18 shows that this plasmid does exhibit the banding pattern expected for the correct pBS/2xAsc/Tc8-5 clone. Fig.19 shows a map of the pBS/2xAsc/Tc8-5 clone.

### **Site Directed Mutagenesis to create a unique *NdeI* Site in the *Protamine-2* coding region**

Once the desired pBS/2xAsc/Tc8-5 clone had been obtained, the next step was to create a unique restriction site at the C-terminus of the *Protamine-2* coding region so that a fluorescent tag could be inserted in-frame, producing a *Protamine-2*-GFP fusion construct. The C-terminus was chosen as the site for inserting the GFP cassette since previous work on *Drosophila* protamines demonstrated that protamines with C-terminal GFP or RFP tags were normally expressed and apparently fully functional (Jayaramaiah-Raja and Renkawitz-Pohl, 2005; Manier, et al., 2010). Because there was an existing plasmid construct containing the GFP coding sequence flanked by *NdeI* sites that could be used as the source of a GFP cassette, and because there was no *NdeI* site already present in the pBS/2xAsc/Tc8-5 clone, I used site-directed mutagenesis to create a unique *NdeI* site at the end of the *Protamine-2* coding sequence to be used for inserting

Fig. 17

Fig. 18

Fig. 19

the tag. Fig. 20 shows the sense/anti sense mutagenesis primer sequence used to create the new *NdeI* site at the end the *Protamine-2* coding region.

After annealing the synthetic oligonucleotide primers to the denatured pBS/2xAsc/Tc8-5 plasmid, the strand-displacing action of *PfuTurbo* DNA polymerase will extend and incorporate the mutagenic primers resulting in nicked circular strands containing the desired mutation. The methylated non-mutated parental DNA template, which became methylated during growth in the host *E. coli* cells, is then digested by *DpnI* treatment, while the *in vitro* synthesized new mutant strands are left intact. After transforming the sample into competent *E. coli*, the resulting colonies were picked, grown up overnight in liquid culture, and subjected to plasmid DNA miniprep procedure. The DNA was then checked by restriction enzyme digest using *NdeI*. Since there should only be one *NdeI* site in the newly mutated pBS/2xAsc/Tc8-5+Nde plasmid, the resulting banding pattern should reveal one band of about 4.2 kb (Fig. 21). As seen, samples # 1, 4, 6, and 7 are all the desired pBS/2xAsc/Tc8-5+Nde constructs.

#### **GFP-tagging of the *Protamine-2* gene in the pBS/2xAsc/Tc5-8+NdeI construct**

The GFP cassette, flanked by *NdeI* cutting sites, is located in a plasmid called pTOPO/GFP/Nde3 (Manier, et al., 2010). This plasmid was treated with *NdeI* (to cut out the GFP coding sequence) and *ScaI* (to cut the pTOPO vector in two, disabling its ability to recircularize without the GFP insert). The pBS/2xAsc/Tc8-5+Nde#1 plasmid was treated with *NdeI* to cut the plasmid at the



Fig. 20

Fig. 21

newly created *NdeI* site within the *Protamine-2* gene sequence. The plasmid was also treated with SAP to dephosphorylate the 5' nucleotides to prevent it from rejoining without the GFP cassette. The resulting bands from pTOPO/GFP/*Nde3* after being cut with *NdeI* and *ScaI* are 2.1 kb and 1.8 kb, representing the *ScaI*-cut pTOPO vector, and 0.9 kb from the GFP *NdeI* cassette. *NdeI* digestion of pBS/2xAsc/Tc8-5+Nde#1 yields one band of 4.2 kb representing the linearized plasmid. Agarose gel electrophoresis of the two digests confirmed that the proper fragments were produced. The DNAs were then mixed and a ligation reaction was performed. Two reactions were set up, one was the ligation reaction using T4 ligase, and the other was a no ligase control reaction. The ligase and no ligase reactions were transformed by heat shock into chemically competent *E. coli* cells. After overnight incubation, multiple colonies were picked and plasmid minipreps carried out. The resulting purified DNA was then initially checked by *EcoRI* digestion. Fig. 22 shows the restriction map of the desired Tc8-5/*Nde*+GFP construct. The GFP cassette has an *EcoRI* site, therefore, the insert is in the proper orientation, *EcoRI* digestion should give three bands: a 2.9 kb band representing the vector, and 1.9 kb and 0.3 kb bands representing the Tc8-5-GFP digested fragments. If the GFP is inserted in the wrong orientation, the three bands will be 2.9, 1.3 and 0.9 kb. If there is no GFP insert, then there will be two bands of 2.9 and 1.3 kb. As seen in Fig. 23, clones #1 and #13 appear to have the GFP insert, with #13 showing the correct orientation and #1 being in the wrong orientation. To confirm that #13 was the correct plasmid construct, two additional digests (e.g., *NdeI* and *AscI*) were carried out. Fig. 24 shows that the correct pattern was

Fig. 22

Fig. 23

Fig. 24

observed. Therefore, this construct, named pBS/2xAsc/Tc8-5-GFP#13 was used for the next step.

### **Insertion of the Tc8-5 GFP-tagged *Protamine-2* gene into the transformation vector pB3xP3-EGFPaf**

After successful creation of the pBS/2xAsc/Tc5-8-GFP construct, the next step before introducing it into the *T. castaneum* genome was to subclone the construct into another plasmid called pB3xP3-EGFPaf (see Fig. 25). This vector carries sequences from an insect transposable element, called *piggybac*, which allows one to introduce the cloned DNA back into the *Tribolium* genome. To do this, the DNA is injected, along with another “helper” plasmid that encodes the *piggybac* transposase enzyme, into the germline region of *Tribolium* embryos. The transposase then catalyzes the transposition of the *piggybac* plasmid construct into the genome. After the injected embryos become adults, they are mated and the offspring scored for presence of the transgene. This vector also carries as a genetic marker a GFP coding region downstream of a strong eye-specific promoter sequence, 3xP3/EGFP. Transgenic beetles are therefore easily recognized by their fluorescent green eyes (Fig. 26).

The pB3xP3-EGFPaf vector has a unique *AscI* site that can be used as a cloning site. So, pBS/2xAsc/Tc8-5-GFP#13 was digested with *AscI* to cut out the insert. Normally, I would also simultaneously treat with *ScaI* to cut up the pBS/2xAsc vector to prevent it from recircularizing. However, because the

Fig. 25



Fig. 26

Tc8-5-GFP sequence contains a *ScaI* site, this was not feasible. The pB3xP3/EGFPaf vector was also cut with *AscI* to linearize it, and treated with SAP to prevent its recircularization. The DNAs were mixed, treated with DNA ligase, and transformed into *E. coli*. Because the pBS/2xAsc vector DNA was present in the reaction mix, a large number of colonies were expected, with most coming from the recircularized pBS/2xAsc plasmid. These however should yield blue colonies when plated on X-gal, and so they could be distinguished from the desired clones which should give white colonies.

Transformation of the ligation reaction produced hundreds of colonies, both blue and white. Fourteen white colonies were picked, grown up in liquid culture overnight, and subjected to the plasmid miniprep procedure. The DNAs were initially cut with *EcoRI* to see if any looked like the desired pB3xP3/EGFPaf/Tc8-5-GFP clone. This digest of a pB3xP3/EGFPaf/Tc8-5-GFP clone is expected to give four bands: 5.7 and 1.6 kb bands representing the cut vector and 1.9 and 0.3 representing the cut Tc8-5-GFP insert. As shown in Fig 27, samples # 1, 2, 6, 9, 10, 12, and 13 showed the correct banding pattern. Samples 10 and 12 were then further analyzed by cutting with *AscI*, *PstI* and *NotI* to confirm that they were correct. The results confirmed that both of these clones were the desired pB3xP3/EGFPaf/Tc8-5-GFP constructs, however they were different in that the inserted fragment was in opposite orientations in the two clones. Fig. 28 shows the restriction maps of these constructs.

Fig. 27

Fig. 28

## CONCLUSIONS AND FUTURE DIRECTIONS

This thesis describes the identification of a cluster of three genes that are candidates for encoding *Tribolium* protamines. In my work, I have used PCR to amplify these genes and have cloned two of them (*Protamine-2* and *Protamine-3*) into a plasmid vector for further analysis. My initial efforts focused on the *Protamine-2* gene. I have used site-directed mutagenesis to create a unique restriction site at the end of its coding region, and have inserted a GFP coding sequence, in-frame, to create a *Protamine-2-GFP* fusion gene. After confirming the identity and proper structure of this construct, I subcloned the fragment into a transformation vector that will allow it to be introduced into the *Tribolium* genome. A very similar set of experiments can now be done to GFP-tag the *Protamine-3* gene, carried on the pCR2.1 TOPO/Tc7-9 clone described above.

The next step, which is beyond the scope of this thesis, will be to carry out germline transformation procedures, using these GFP-tagged protamine candidate genes to create transgenic beetles. Examination of the developing sperm will then be done to (1) see if the sperm heads are labeled with GFP (which will provide strong support for the hypothesis that this is a protamine gene) and (2) see if the timing of the histone-protamine transition is conserved between *Drosophila* and *Tribolium*.

That these experiments will likely yield useful results is suggested by the experiments done by Dr. Belote with the *Protamine-1* gene, amplified using the Tc2-4 primer set described above. In that case, a GFP-tagged *Protamine-1* was

constructed and cloned into pB3xP3/EGFPaf using the same methods I describe here for *Protamine-2*. Transgenic beetles were then generated and sperm examined for expression of the Protamine-1-GFP. As shown in Fig. 29, the sperm heads were labeled, providing strong evidence that this is indeed a protamine gene. Thus, it is highly likely that the similar *Protamine-2* and *Protamine-3* genes are also *bona fide* protamine genes. It will be of interest to see if there are any differences in the expression of these three genes. Ultimately it would be interesting to know whether these genes are functionally redundant, or if they are somehow different.

Fig. 29

**REFERENCES**

- Balhorn R (2007) The protamine family of sperm nuclear proteins. *Genome Biology*. **8**:227.
- Berghammer A, Bucher G, Maderspacher F, Klinger M (1999) A system to efficiently maintain embryonic lethal mutations in the flour beetle *Tribolium castaneum*. *Dev Genes Evol*. **209**: 382-389.
- Bloch-Qazi, MC (2003) A potential mechanism for cryptic female choice in a flour beetle. *Journal of Evolutionary Biology*. **16**: 170-176.
- Brooker RJ, (2009) *Genetics: Analysis and Principals*. 3<sup>rd</sup> Edition. McGrawHill. Boston, MA.
- FlyBase Homepage. Web. 28 Apr. 2011. <http://flybase.org>.
- Fuller MT (1998) Genetic control of cell proliferation and differentiation in *Drosophila* spermatogenesis. *Seminars in Cell & Developmental Biology*. **9**: 433-44.
- Gaucher J, Reynoird N, Montellier E, Boussouar F, Rousseaux S, Khochbin S. (2010) From meiosis to postmeiotic events: The secrets of histone disappearance. *The FEBS Journal* **277**: 599-604.



Henikoff, S (2008) Nucleosome destabilization in the epigenetic regulation of gene expression. *Nature Review*. **9**: 15-22.

Horn C, Wimmer EA (2000) A versatile vector set for animal transgenesis. *Developmental Genes and Evolution*. **210**: 630-637.

Jayaramaiah S, Raja SJ, Renkawitz-Pohl R (2005) Replacement by *Drosophila melanogaster* Protamines and Mst77F of Histones during Chromatin Condensation in Late Spermatids and Role of Sesame in the Removal of These Proteins from the Male Pronucleus. *Molecular and Cellular Biology*. **25**: 6165-6177.

Lewis JD, Abbott DW, Ausio J (2003) A haploid affair: core histone transitions during spermatogenesis. *Biochem Cell Biology*. **81**:131-40.

Lorenzen MD, Berghammer S, Brown J, Denell RE, Klinger M, Beeman RW (2003) *piggyBac*-mediated germline transformation in the beetle *Tribolium castaneum*. *Insect Molecular Biology*. **12**: 433-440.

Olivia R, Castillo J (2011) Proteomics and the genetics of sperm chromatin condensation. *Asian Journal of Andrology*. **13**: 24-30.

Manier MK, Belote JM, Berben KS, Novikov D, Stuart WT, Pitnick S (2010) Resolving Mechanisms of Competitive Fertilization Success in *Drosophila melanogaster*. *Science*. **328**: 354-357.

Michalczyk L, Martin OY, Millard AL, Emerson BC, Gage MJG (2010) Inbreeding depresses sperm competitiveness, but not fertilization or mating success in male *Tribolium castaneum*. *Proceedings of the Royal Society: Biological Sciences*. **10**: 1-9.

Muller HJ, Altenburg E (1919) The rate of change of hereditary factors in *Drosophila*. *Proceedings of the Society of Experimental Biology and Medicine*. **17**: 10-14.

Rathke C, Baarends WM, Jayaramaiah-Raja S, Bartkuhn M, Renkawitz R, Renkawitz-Pohl R (2007) Transition from a nucleosome-based to a protamine-based chromatin configuration during spermiogenesis in *Drosophila*. *Journal of Cell Science*. **120**:1689-1700.

Rathke C, Barckmann B, Burkhard S, Jayaramaiah-Raja S, Roote J, Renkawitz-Pohl R. (2010) Distinct functions of Mst77 and protamines in nuclear shaping and chromatin condensation during *Drosophila* spermatogenesis. *European Journal of Cell Biology*. **89**: 321-338.

Retief JD, Dixon GH (1993) Evolution of pro-protamine P2 genes in primates.  
The FEBS Journal. **10**: 609-615.

Sokoloff, A (1974) The Biology of Tribolium. Oxford University Press. Oxford.

Stored Product Insect Research Unit : Tribolium Stock Maintenance. *ARS : Home*.

Web. 06 May 2011.

<<http://www.ars.usda.gov/Research/docs.htm?docid=12892>>.

## WRITTEN CAPSTONE SUMMARY

This thesis describes the identification of a cluster of three genes that are candidates for encoding *Tribolium* protamines. Protamine genes are small basic proteins, similar to histones, which work to condense DNA during spermatogenesis. In my work, I have used PCR, Polymerase Chain Reaction, to amplify these genes. PCR is a useful tool that can be used in order to target and amplify a certain piece of genomic DNA using primers that flank the desired sequence. I have cloned two of the *Tribolium* protamine genes (*Protamine-2* and *Protamine-3*) into a plasmid vector for further analysis. A plasmid vector is a small circular piece of DNA that is used commonly as a means to manipulate genes. My initial efforts focused on the *Protamine-2* gene. I have used site-directed mutagenesis to create a unique restriction site at the end of its coding region, and have inserted a Green Fluorescent Protein (GFP) coding sequence, in-frame, to create a *Protamine-2-GFP* fusion gene. GFP originally comes from a jellyfish, and allows for the fluorescent visualization of a gene in which it is inserted into. After confirming the identity and proper structure of this construct, I subcloned the fragment into a transformation vector that will allow it to be introduced into the *Tribolium* genome. The transformation vector is injected into the early *Tribolium* embryos. A very similar set of experiments can now be done to GFP-tag the *Protamine-3* gene, carried on the pCR2.1 TOPO/Tc7-9 clone described above.

Beyond the scope of this thesis, would be to carry out germline transformation procedures, using these GFP-tagged protamine candidate genes to

create transgenic beetles. Examination of the developing sperm will then be done to (1) see if the sperm heads are labeled with GFP (which will provide strong support for the hypothesis that this is a protamine gene) and (2) see if the timing of the histone-protamine transition is conserved between *Drosophila* and *Tribolium*.

That these experiments will likely yield useful results is suggested by the experiments done by Dr. Belote with the *Protamine-1* gene, amplified using the Tc2-4 primer set described above. In that case, a GFP-tagged Protamine-1 was constructed and cloned into pB3xP3/EGFPaf using the same methods I describe here for *Protamine-2*. Transgenic beetles were then generated and sperm examined for expression of the Protamine-1-GFP. The sperm heads were labeled, providing strong evidence that this is indeed a protamine gene. Thus, it is highly likely that the similar *Protamine-2* and *Protamine-3* genes are also *bona fide* protamine genes. It will be of interest to see if there are any differences in the expression of these three genes. Ultimately it would be interesting to know whether these genes are functionally redundant, or if they are somehow different.