May 2006

Regulation of E(spl) Gene Expression During Development

Morgan L. Maeder
Connecticut College, mlmae@conncoll.edu

REGULATION OF *E(spl)* GENE EXPRESSION DURING DEVELOPMENT

by
Morgan Lee Maeder

A thesis submitted in partial fulfillment
of the requirements for the degree of

Bachelor of Arts

in Biology, Cell and Molecular Concentration

Connecticut College

2006

Approved by _____
Chairperson of Thesis Committee

Date _____

Thesis Committee:

Professor Deborah Eastman, Ph.D., Advisor
Department of Biology, Connecticut College

Professor Anne Bernhard, Ph.D.
Department of Biology, Connecticut College

Professor Bruce Branchini, Ph.D
Department of Chemistry, Connecticut College

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to thank Professor Deborah Eastman for her unending guidance and support. This year has been an invaluable and unparalleled experience. She has been both a role model and a friend, and her enthusiasm has driven me to become even more of a science dork than I thought possible.

I would also like to thank Beto Zuniga for all of his help with absolutely everything. There are many days that I would not have survived without Beto's extensive knowledge and patience, and just as importantly his sense of humor.

Thank you to the many people who took the time to read my thesis and offer me their thoughts – Professors Anne Bernhard, Bruce Branchini, Phillip Barnes, and of course Deborah Eastman, who has read this more times than I can count. Thank you also to Anne Bernhard for all her help with the Real Time PCR – for her generous donation of materials and her time. Thank you to Tara Southworth for her help with the luciferase assay and to Page Owen for his help with larva photos.

The friendship and support of my lab mates, Jeff Bender, Katey Nelson, Ben Polansky and Bryanne Robson has been invaluable to me, both in and out of the lab. With them I have shared moments of craziness, frustration and excitement over results, and late nights in the lab. I can always count on them to make organic chemistry mom jokes and get excited about colored eppendorf tubes and what strawberries look like under a microscope. Their presence has made New London Hall a fun place to spend my time.

ABSTRACT

REGULATION OF *E(spl)* GENE EXPRESSION DURING DEVELOPMENT

By Morgan L. Maeder

Chairperson of the Thesis Committee: Dr. Deborah Eastman
Department of Biology

The Notch pathway, a crucial developmental signaling system, acts to direct the fates of individual cells in many organisms and has also been implicated in a wide range of human diseases. Notch signaling plays a vital role in cell fate decisions in almost every tissue type ranging from the skin to the nervous and vascular systems. Aberrant Notch signaling has been implicated as a cause of many diseases, including a variety of cancers. Activation of the Notch receptor releases a Notch intracellular domain into the nucleus, where it binds with a transcription factor, Suppressor of Hairless (Su(H))to create an active complex which upregulates expression of target genes. In *Drosophila* the primary targets of Notch activation are the *Enhancer of Split [E(spl)]* genes. The *E(spl)* genes encode a family of basic-helix-loop-helix (bHLH) transcription factors, which exhibit overlapping functions throughout developmental stages. In order to determine the mechanisms through which *E(spl)* gene expression is controlled, I used three approaches to study *E(spl)* regulation. First, Bioinformatics analysis of the upstream regulatory regions of the *E(spl)* genes reveals binding sites for transcription factors that may act to regulate *E(spl)* gene expression. Evolutionary conservation of sites in the regulatory region lends support to their importance in the regulation of gene expression. Second, Real Time PCR quantification of the expression of three *E(spl)* genes at different stages of *Drosophila* metamorphosis suggest roles for some of these genes. Third, a reporter vector with the upstream region of one of the *E(spl)* genes cloned upstream of the firefly *luciferase* gene was

constructed and used in *Drosophila* tissue culture experiments to further analyze the regulation of gene expression. Results from these three approaches will help to better understand the process of gene regulation and to characterize the mechanisms involved in controlling gene expression. Specific understanding of Notch target genes will elucidate how the Notch pathway functions in both normal and disease cells.

Introduction

The development of a complex multicellular organism from a single fertilized egg is the result of the intricate coordination of differentiation, cell proliferation, growth, and programmed cell death. This process involves translating genetic information, inherited from parents, into the physical traits that will characterize every cell within each individual offspring. Genetic information, encoded in DNA, is manifested as physical characteristics through the production of RNA and protein products. Proteins are involved in nearly all functions of the cell and provide the basis for differences between cell types. The production of proteins, therefore, is one of the most important, and most highly regulated cellular processes. DNA is transcribed into messenger RNA (mRNA) in the nucleus and subsequently translocated into the cytoplasm and translated into proteins. The level of protein production may be controlled through regulation of mRNA transcription from DNA. Transcription is controlled by complex systems of activators and repressors, which respond to extracellular signals and bind to transcription factor binding sites, specific DNA sequences upstream of the gene. The binding of these transcription factors to the upstream regulatory region acts either to upregulate or downregulate transcription of a specific DNA segment and, consequently, affects the expression of the gene into a protein product.

Many signaling pathways within the organism mediate regulatory processes, such as the regulation of gene expression, during development. The Notch pathway is one of these developmental signaling systems. It is a highly conserved pathway that directs the developmental fates of individual cells and is crucial for establishing cell fate distinctions between neighboring cells. In addition to its role during development, Notch has been implicated in many different diseases, such as schizophrenia, Alzheimer's disease and cancer.

Linkage disequilibrium mapping of the human major histocompatibility complex region in 80 sets of parent-offspring trios in which the offspring was affected with schizophrenia, showed that *NOTCH4* is highly associated with schizophrenia (Wei and Hemmings, 2000). Notch disfunction in the adult central nervous system of *Drosophila* has been shown to impair long-term memory by affecting structural neuroplasticity. Mutations in the *Presenilin* gene, an important component of the Notch pathway, have been associated with early onset Alzheimer's disease (Presente et al., 2004). Notch has also been implicated in a variety of cancer types. Notch signaling is crucial for normal cell proliferation and differentiation and when the pathway does not function properly, it often leads to the development of tumors. One study found that induced inactivation of Notch1 in the prostate led to enhanced proliferation of epithelial cells and that Notch1 was downregulated in human prostate adenocarcinomas, indicating that deactivation of Notch signaling plays a role in the development of prostate cancer (Wang et al., 2006). Another study found that Notch1 signaling was upregulated in melanoma cell lines, and indicated a stage-specific role for Notch signaling in advancing the progression of primary melanomas (Balint et al., 2005). While the connection between Notch signaling and the development of these, and many other diseases, is not fully understood, it is clear that Notch affects a wide array of cellular processes in many species, including humans. A deeper understanding of the Notch system could allow for the potential development of treatments for many of the major diseases affecting humans today.

Cell-Cell Signaling Through the Notch Pathway

The Notch pathway has been studied in many organisms, including *Drosophila*, roundworms, sea urchins, frogs, fish, chickens, mice and humans (reviewed in Schweisguth,

2004). In both vertebrate and invertebrate systems, Notch has been shown to play a major role during development. There are three basic mechanisms through which the Notch pathway establishes cell fate decisions: lateral inhibition, binary cell fate and inductive communication. Lateral inhibition is a process in which a cell inhibits nearby cells from adopting its own fate; binary cell fate decisions endow different properties on sister cells as a result of asymmetric cell division; and inductive communication between rows of cells leads each row to adopt a distinct fate (Castroet al., 2005; Buescher et al., 1998; Wakamatsu et al., 2000). Studies in *Drosophila*, in which the embryonic tissues of organisms containing Notch mutations were labeled, showed that the loss of Notch signaling results in abnormalities in many different organs derived from all three germ layers (Hartenstein et al., 1992). The Notch pathway has also been shown to play a crucial role in the maintenance of neural stem cells (Hitoshi et al., 2002). It is necessary for the transition from the "primitive" to the "definitive" stem cell and for the maintenance of the definitive neural stem cell state (Hitoshi et al., 2004).

Notch Pathway Components

The Notch pathway allows for communication between two nearby cells (Figure 1). The signaling cell possesses a ligand, in the case of *Drosophila melanogaster* either Delta or Serrate, while the receiving cell has a transmembrane Notch receptor. When the ligand binds to the receptor, it induces proteolytic cleavage, which results in the release of the Notch intra-cellular domain (NICD) into the cytoplasm of the receiving cell. NICD is the active form of the Notch receptor, and it is translocated into the nucleus of the cell, where it acts as a transcriptional regulator. In the nucleus NICD associates with a transcription factor to activate transcription of target genes. In *Drosophila melanogaster*, this transcription factor is Suppressor of Hairless

[Su(H)]. Prior to Notch activation, Su(H) is already bound to a specific DNA sequence in the upstream regulatory region. However, it is bound to co-repressors and therefore acts as a repressor. NICD disrupts the co-repressors bound to Su(H) and replaces them. This NICD/Su(H) complex is then able to activate gene expression. In *Drosophila* the primary targets of Notch are the *Enhancer of Split [E(spl)]* genes (Castro et al, 2005). Cell-specific activation of certain Notch target genes requires interaction between the NICD and specific proneural activator proteins (Cooper et al, 2000). Transcriptional synergy between the NICD and proneural activator proteins requires a specific organization of the upstream regulatory region of the target genes. Experiments have shown that a pair of Su(H) binding sites, that are specifically oriented, are necessary for Notch-proneural function, wherein this DNA architecture allows for synergistic interactions with proneural proteins (Cave et al, 2005). Notch acts to upregulate the seven *E(spl)* genes and thus increase production of their protein products (Jennings et al, 1994, Eastman et al, 1997).

Drosophila **Neurogenesis**

The Notch signaling pathway has been most extensively studied in the context of fly neurogenesis. In the peripheral nervous system, each sensory organ of the adult fly is derived from a single sensory organ precursor cell (SOP). The SOPs arise out of clusters of cells, known as proneural clusters (PNCs), present in the imaginal discs. Imaginal discs are foldings of undifferentiated epithelial sheets in the larva, which will eventually give rise to the physiological structures of the adult fly during metamorphosis. Determination of the SOPs occurs during the late larval and early pupal stages of metamorphosis when the genes *achaete* and *scute* confer upon one cell in the PNC the ability to become an SOP. The *achaete* and *scute* genes encode a

set of basic Helix-Loop-Helix proneural proteins, which bind to enhancer sites, called E boxes. These proteins regulate the expression of neuronal specific genes and are necessary for a cell to take on a neuronal fate. The remaining cells of the proneural cluster are inhibited from becoming neuronal cells by lateral inhibition and instead are relegated to an epidermal fate (Culi and Modolell, 1998). Lateral inhibition of the surrounding cells occurs via Notch-mediated signaling (Hartenstein and Posakony, 1990; Schweisguth and Posakony, 1994). A similar process occurs in the ventral neuroectoderm of the early fly embryo and establishes the central nervous system (for review, Artavanis-Tsakonas et al. 2003).

Enhancer of Split **Genes**

The *E(spl)* locus contains 12 different genes, seven of which encode a family of basic-helix-loop-helix (bHLH) transcription factors (Figure 2). The expression of the *E(spl) bHLH* genes is transcriptionally upregulated by activated Notch (Cooper et al, 2000). The seven bHLH proteins act as transcriptional repressors and bind to specific sites, known as N boxes. These proteins exhibit overlapping functions throughout developmental stages. However, it is not yet completely understood why such a redundancy exists (Delidakis and Artavanis-Tsakonas, 1992). It has been shown that individual *E(spl)* proteins are more effective than others at influencing certain processes (Ligoxygakis et al, 1999). For example *E(spl)mγ* is the most powerful of the seven bHLH proteins in the wing imaginal disc, but it is less potent in other tissues. In the wing imaginal disk, *mγ* acts as a transcriptional repressor to suppress the proneural genes *achaete* and *scute* as well as multiple other targets which are not yet thoroughly understood. Ectopic and over-expression of the E(spl) proteins in imaginal disks showed *mβ* to produce a dramatic loss of

wing veins, while *mγ* was most potent in repressing the wing pouch and in suppression of wing nicking (Ligoxygakis et al., 1999).

The *E(spl)* genes show unique, but overlapping expression patterns in the D. melanogaster embryo (Wei and Hemmings, 2000). In larval imaginal discs they show more distinct expression patterns (de Celis, 1996). *E(spl)mδ* is transcribed in all sensory organ clusters, while *mδ* and *mγ* are transcribed in a only a subset of sensory clusters, but strongly in the developing eye. *E(spl)mβ* is transcribed heavily in the intervein regions and the dorsal/ventral boundaries of the wing and eye imaginal discs. The distinct expression of these *E(spl)* genes, combined with the finding that activated Notch and Suppressor of Hairless are capable of eliciting only limited transcription of the genes (Cooper, and Tyler, et al, 2000), indicates a requirement for additional transcriptional regulators that direct expression of the *E(spl)* genes.

Although *mγ* is upregulated in response to proneural proteins and NICD, it failed to respond to these activators in the wing pouch (Cooper et al, 2000). A vector construct containing 10 Su(H) high affinity binding sites cloned upstream of the *mγ* promoter region was injected into fly embryos, and transformants showed a gene expression pattern similar to endogenous *mγ* activity (Go et al. 1998). The expression patterns of *mγ* suggest the presence of a repressor that acts to prevent expression in the wing discs. Examination of the upstream regulatory region of *mγ* showed the presence of three Tramtrack69 (Ttk69) binding sites in *D. melanogaster* and two Ttk69 sites in *D. hydei* (Figure 3, personal communication with D. Eastman). Tramtrack is a zinc finger transcription factor that is known to play an essential role in multiple aspects of development. The conserved presence of Ttk69 binding sites in the regulatory region of *mγ* suggests that it may act in combination with Notch to regulate *mγ* expression.

Tramtrack 69 as a Possible Repressor of *E(spl)mγ*

The *tramtrack (ttk)* gene encodes two alternatively spliced DNA-binding proteins – Ttk69 and Ttk88. Ttk69 is known to repress various pair-rule genes, which help to establish the body plan in *Drosophila* embryos. Experiments have shown that the repressive ability of Ttk69 does not always require the binding of Ttk69 to DNA. A part of Ttk69, the POZ/BTB domain, does not by itself bind DNA. However, while it is not as efficient as the full-length Ttk69, it was found capable of repressing GAGA-dependent activation of a pair-rule gene promoter (Pagans et al., 2004). Ttk69 is expressed outside of the proneural clusters where it acts as a blanket repressor, as well as in sensory organ precursor daughter cells where it inhibits transcription of *achaete* and *scute* to prevent a neuronal fate (Badenhorst et al., 2002). In addition to downregulating the *achaete* and *scute* genes, it appears that Tramtrack may also act to repress *E(spl)mγ* in the wing imaginal disks. Because *mγ* represses transcription of the *achaete-scute* complex, this gives Ttk the ability both to upregulate and to downregulate the expression of the *achaete* and *scute* proneural genes. This could be a mechanism for ensuring the tight regulation of these genes because it is imperative that the proneural genes are present only in precise amounts. Ttk's potential ability to suppress *mγ* could result from a similar need to tightly control the expression of this E(spl) protein in the wing imaginal disc, where it is the most potent of the seven *E(spl)* gene products (Ligoxygakis et al., 1999**)**. Preliminary evidence using reporter assays in tissue culture cells and overexpression studies in imaginal discs suggest that Ttk69 may indeed repress *mγ* expression (D. Eastman, unpublished results).

A Role for Notch Signaling in *Drosophila* Gut Development

The Notch pathway has also been shown to play a role in *Drosophila* gut development. Notch signaling in the hindgut controls the fate of a single row of boundary cells separating the dorsal and ventral halves of the gut tube (Fusse and Hoch, 2002). The *Drosophila* gut consists of three regions; the foregut, midgut and hindgut. The development of the gut begins during gastrulation with the invagination of ectodermal cells to give rise to the foregut and hindgut primordial tubes. The midgut is formed in between these tubes and is derived from endodermal cells that migrate from both types of primordial tubes. This primitive gut tube is then remodeled during metamorphosis and divided into the three distinct regions. The developing hindgut becomes divided along anterior-posterior and dorsal-ventral axes. The three regions of the gut are then further developed into distinct organs (For a review of gut development, Lengyel and Iwaki, 2002). The steroid hormone 20-hydroxyecdysone (ecdysone) is active in the development of the gut during metamorphosis. The two major types of cells in the larval midgut both respond, in different ways, to ecdysone. Pulses of ecdysone direct the destruction of obsolete larval epidermal cells through programmed cell death and initiate their replacement by adult epidermal progenitor cells, which will eventually form the tissues of the adult fly (Jiang et al., 1997; Lee et al., 2002).

Ecdysone regulates changes in gene expression during metamorphosis and elicits different morphological and physiological responses in different tissues throughout development (Beckstead et al., 2005; Fletcher et al., 1995; Schubiger et al., 1998). Ecdysone binds directly to an ecdysone receptor, which stimulates the active transcription of DNA. The binding of ecdysone acts to mediate polytene chromosome puffing, which is a visible indication of active transcription. The ecdysone-inducible *E74* gene has been shown to play a critical role in

Drosophila metamorphosis. Loss-of-function analysis has shown that mutations in both transcription units of the *E74* gene, *E74A* and *E74B*, are predominantly lethal, suggesting that *E74* is required for pupation and for the metamorphosis of both larval and imaginal tissues (Fletcher et al, 1995). Studies of gene expression in flies containing a mutant ecdysone receptor allele showed that 76% of genes that are significantly induced at the initiation of midgut metamorphosis in *Drosophila* require a functional ecdysone receptor and are therefore under the control of transcription factors mediated by ecdysone (Li, and White, 2003). These results also implicate a connection between the ecdysone regulatory network and the Notch signaling pathway. The results of microarray techniques, which measure the extent of gene activation, showed that *E(spl)* genes, specifically *mα* and *mβ*, are, at least in part, induced by ecdysone signaling (Li and White, 2003).

In silico Approaches to the Study of Gene Regulation

In silico Biology is a relatively new field of science which combines biology, computer science and information technology. Born out of major advances in molecular biology and genomic technology over the past few decades, the explosion of biological information has made it necessary to establish databases for the storage and analysis of these data. *In silico* biology, also called Bioinformatics, involves the development of tools to organize this information and the creation of algorithms to analyze it. Over the past year, the complete genomes of 9 different Drosophila species have been sequenced and annotated*; D. simulans, D. yakuba, D. erecta, D. ananassae, D. pseudoobscura, D. persimilis, D. virilis, D. mojavensis* and *D. grimshawi*. In addition, several new bioinformatics tools have been developed and are now available. Specifically, BLAT, similar to BLAST, is a whole-genome search algorithm that contains an

index for the genomes of numerous species (Kent, 2002). BLAT works by generating a comparison of the reference and test-species DNA sequence alignments, highlighting the specific bases that differ between the two DNA sequences. EVOPRINTER is a database used for the identification of evolutionarily conserved DNA sequences. EVOPRINTER is a multispecies-alignment program which uses BLATs to generate an evolutionary gene print (*EvoP*) displaying the reference DNA sequence and highlighting all of the multispecies-conserved DNA sequences (Odenwald, and Rasband, et al, 2005). EVOPRINTER makes it possible to observe the conservation of binding sites in the DNA sequences for regulatory elements (such as Su(H) binding sites, E boxes and N boxes) across the different *Drosophila* species. Other DNA binding protein sites that are conserved across species can be identified using MatInspector. MatInspector is an algorithm that utilizes nucleotide and position weight matrices to identify potential transcription factor binding sites in the DNA (Cartharius et al., 2005). This program quickly scans an input genetic sequence and generates a list of all identified binding sites and their location in the DNA sequence. The evolutionary conservation of genetic sequences indicates that these sequences have an important function, and the conserved binding sites identified by MatInspector can therefore be expected to be crucial to *Drosophila* development. Examining the conservation of these DNA sequences across species may also provide phylogenetic information for the *Drosophila* genes throughout evolution.

Approaches to Examine Regulation of *E(spl)* Gene Expression

Although it is known that Notch signaling plays a vital role in *E(spl)* expression, it is not yet understood what other mechanisms are active in *E(spl)* regulation. Most likely, a myriad of transcription factors interact to control *E(spl)* expression, however, the identity of these factors

remains unknown. In my thesis I seek to provide insight into the functioning of the Notch pathway by using three different approaches to examine the regulation of *E(spl)* gene expression during development. Bioinformatics analysis was used to examine the upstream regulatory regions of the *E(spl)* genes. Examination of the evolutionary conservation of sites known to play a role in *E(spl)* gene regulation provides further evidence for the importance of these sites. The regulatory regions were scanned for transcription factor binding sites. The presence of a transcription factor binding site could implicate that specific factor in the regulation of *E(spl)* gene expression. Bioinformatics analysis showed a high level of conservation of Su(H) sites, E boxes and N boxes across *Drosophila* species. The conservation of these sites strongly implicates in the regulation of *E(spl)* gene expression. The Notch pathway and Ttk69 are crucial for many stages of development, and preliminary experiments suggest that their interaction may be a mechanism for regulation of *mγ* expression. An *mγ* luciferase reporter vector was constructed, and preliminary evidence shows that it is functional in S2 cell transfection experiments. This vector will be useful in the study of Ttk regulation of *E(spl) mγ* expression. I also examine other factors that may play a role in transcriptional regulation of the *E(spl)* genes. Bioinformatics analysis of the regulatory regions of these genes identified binding sites for possible transcriptional regulators. These were then further tested by Real-Time RT-PCR analysis of *E(spl)* expression during development. Molecular and Bioinformatics analysis has allowed me to examine the regulation of *E(spl)* gene expression in *Drosophila*. My results of the regulation of *E(spl)* genes have contributed to the understanding of the Notch signaling pathway. Knowledge of the Notch pathway is crucial, not only to comprehension of basic developmental processes, but also to the further characterization of diseases which result from abnormal Notch signaling.

Materials and Methods

Bioinformatics

Reference DNA sequences for *Drosophila melanogaster* were obtained from the University of California, Santa Cruz Genome Browser database (http://genome.ucsc.edu/cgi-bin/hgGateway). A particular reference sequence was then copied and pasted into the BLAT search engine window (http://genome.ucsc.edu/cgi-bin/hgBlat) and individually compared to the nine different test species: *simulans, yakuba, erecta, ananassae, persimilis, pseudoobscura, virilis, mojavensis* and *grimshawi*.

The highest-scoring readout alignment for each test species was selected and pasted into an EVOPRINTER input window (http://evoprinter.ninds.nih.gov). EvoPs were generated using subsets of the BLAT inputs as well as BLAT readouts from all of the test species. The EvoPs were saved as a word document and sites of interest (Suppressor of hairless binding sites, N boxes, E boxes, ttk69 binding sites and TATA box) were highlighted. Based on the Drosophila phylogenetic tree, species were sequentially added into Evoprinter in the following order: *D. simulans, D. yakuba, D. erecta, D. ananassae, D. persimilis, D. pseudoobscura, D. virilis, D. mojavensis, D. grimshawi*.

Highly conserved sequences were also highlighted. Potential transcription-factor DNA-binding site functions of these sequences were identified by pasting the entire gene sequence into MATINSPECTOR (http://www.genomatix.de/cgi-bin/matinspector_prof/mat_fam.pl).

BLATs were saved in Microsoft Word, and sequences of interest were identified using the "find" tool. These sequences were compared in the ten different *Drosophila* species. For each of the test species all of the Su(H) sites, N Boxes, E Boxes and ecdysone sites in each gene

were categorized as fully conserved, partially conserved, or not conserved, with reference to *Drosophila melanogaster*.

Construction of mγ Luciferase Reporter Vectors

An attempt was made to insert a small, 280 bp DNA fragment of the *mγ* promoter into the Promega Luciferase vector, pGL3 (Figure 4). Briefly, pGL3 vector was digested with *HindIII* and *XhoI* at 37°C for 2 hours. The 4.8Kb DNA fragment was isolated by gel electrophoresis with low-melt agarose. The DNA sample (30µl) was run on a 1.2% agarose gel with TAE buffer. A section of the gel directly below the band of the large fragment of the vector was removed and filled with 1.2% low-melt agarose and the band was allowed to run into this section. The low-melt agarose section containing the DNA fragment was then removed and incubated at 65°C with 20mM Tris•Cl, 1mM EDTA to melt the gel. The DNA was then purified by phenol-chloroform extraction followed by ethanol precipitation (Sambrook et al, 1989). Vector DNA was resuspended in 20µl of water and the purity of the sample was ascertained by using a spectrophotometer to find the OD_{260}/OD_{280} ratio.

The *mγ* insert was digested out of the *mγ* promoter PCRII vector (constructed by B. Siegel, 2005) using *XhoI* and *HindIII* restriction enzymes to produce a 280 bp DNA sequence. This was isolated using the low-melt gel electrophoresis procedure. Ligation of the *mγ* sequence into the pGL3 vector was attempted using a 1:1 ratio of vector to insert in a reaction mixture containing 0.5 µl of DNA ligase. Ligation reactions were incubated overnight at 16°C.

Due to cloning difficulties with the 280 bp *mγ* promoter fragment, a luciferase vector containing 1.2 kb of the *mγ* promoter was constructed (performed by Beto Zuniga and Deborah Eastman). pGL3 vector was digested with *SmaI*, a restriction enzyme which cuts DNA to

produce blunt ends. To prevent the vector ends from re-ligating together, the digested vector was treated with Antarctic Phosphatase. H2.1, a vector containing the entire *mgamma* sequence, was digested with HindIII and BanII to produce a 1.2Kb fragment containing the *mgamma* upstream regulatory region. A Klenow reaction was done to fill in the 5' overhangs and produce blunt ends on the *mγ* insert. Ligation was performed with1 µl of ligase at a concentration of200,000U/µl, and a 4:1 ratio of insert to vector.

Transformation of Ligation reactions into Competent Cells

MAX Efficiency Stb12 competent bacterial cells were transformed with ligation reactions as described in the Invitrogen protocol. Cells (50 µl) were placed into pre-chilled test tubes and 1µl, of each ligation reaction was added to cell aliquots. The solutions sat on ice for 30 minutes and were then heat shocked at 42°C for 25 seconds and returned to ice for 2 minutes. Room temperature SOC (0.9 ml) was added to each tube, and they were shaken at 225 rpm for 90 minutes at 30°C. The solutions were plated onto Luria-Broth with Ampicilin (LB/AMP) plates at two different concentrations – either 100 µl of solution or 30µl solution + 70µl SOC. The plates were incubated overnight at 30°C.

Mini-Preparations of Ligated Vectors

Tubes containing 5ml of LB-Amp, were inoculated with a bacterial colony picked off of the transformation plates. These cultures were shaken overnight at 37°C. Liquid culture (1.5 ml) from each tube was placed into a sterile tube and centrifuged for 1 minute at max rpm. The supernatant was removed and the pellet was resuspended in 300 µl STET . Lysozyme solution (25µl) was added and mixed and the tubes were immediately placed in boiling water for 45

seconds. The solutions were then centrifuged at 12,000g for 30 minutes and the pellet was removed and discarded. Isopropanol (650µl) was added and mixed and then centrifuged for 10 minutes. The supernatant was poured off and the pellet was allowed to air dry. The pellet was washed with 500µl cold 70% EtOH and centrifuged for 10 minutes at 4°C. The supernatant was removed and the pellet was air dried and then resuspended in 50µl of water.

Confirmation of mγ Insertion

Polymerase Chain Reaction (PCR) , a procedure for the amplification of specific DNA sequences, was performed with *mγ* primers to confirm insertion of the sequence into the vector. PCR total reaction volumes were 15µl and each contained approximately 0.1µg of miniprep sample DNA. PCR was carried out with a denaturation temperature of 95°C, an annealing temperature of 61°C and an elongation temperature of 75°C. Thirty replication cycles were carried out: 45seconds denaturation, 30 seconds annealing and 90 seconds elongation. PCR products were analyzed by gel electrophoresis on a 1.2% agarose gel at 120V.

Correct placement of *mγ* insert in the pGL3 vector was confirmed by DNA sequencing. Sequencing was conducted by the Harvard DNA sequencing facility.

Transfection of *Drosophila* S2 Cells

Drosophila S2 cells were maintained at 25˚C. Cells were transfected with 2 different combinations of expression vectors; negative control and NICD. NICD contains the activated form of Notch cloned into RaHa3, a vector containing the *Drosophila* metalothionine promoter, which is inducible by copper sulfate. All samples contained the *mγ* promoter luciferase vector. Four ml of cells in complete media, containing 12.5% fetal calf serum, were transferred to each

well of two 6-well plates and left in the incubator to adhere to plates. Expression vectors were added to 1.5ml of incomplete media, which lacks fetal calf serum, as follows:

	Luciferase	NICD	RaHa
Neg. Control	3µg		9µg
NICD	3µg	3µg	6µg

Incomplete media (6ml) was combined with 600µl of lipofectin. This solution was added to each of the four test tubes (1.5ml into each tube). The cells were washed three times and 1ml of the DNA:lipofectin solution was added to each well to produce 3 replicates of each sample. Cells were incubated for 6 hours, after which the incomplete media was removed and replaced with complete media. After 17 hours, 3µl of copper sulfate solution was added to each well and swirled. Approximately 24 hours later cells were washed 3 times with 1X PBS and then lysed by swirling for 15 minutes with 500µl of 1X *Firefly* Luciferase Lysis Buffer.

Luciferase Assay

D-luciferin was dissolved in Firefly Luciferase Assay Buffer to produce a concentration of 0.2mg/ml. Twenty microliters of each cell solution sample were pipetted into a well. The Luciferase Assay was run on a Veritas Luminometer. Background luminescence was read for 8 seconds. Firefly Luciferase Assay Solution (100 µl) was added, by the veritas machine, to each sample well and luminescence was read again for 8 seconds. Protein analysis was performed on samples by the Bradford Protein Assay from BioRad, to determine protein concentration of the solution. Background luminescence was found to be negligible and relative luminescence units were graphed as a function of ng of protein in the sample.

Staging of *Drosophila* larvae and prepupae

In order to determine the developmental stage of flies, *Drosophila melanogaster* stocks were raised on food supplemented with 0.05% bromophenol blue. Consumption of the food by larvae results in blue guts, which clear as the larvae proceed through development and stops eating (Figure 5). Wandering and stationary third instar larvae were collected and staged based on the amount of bromophenol blue in their intestines. Wandering larvae with dark blue intestines were staged as early third instar larvae, approximately 12-24 hours prepupariation. Those with partially clear guts will pupariate in 5-12 hour and those with completely clear guts will pupariate in 1-6 hours (Andres and Thummel, 1994). Individuals at puparation were obtained by selecting animals whose cuticle had not yet tanned, as tanning of the cuticle is first evident 15-30 minutes after puparium formation (Andres and Thummel, 1994). Some of these newly pupariated animals were allowed to develop an additional 2-3 hours.

Isolation of Midgut Tissue

Larvae and pupae at the correct stages were selected and placed into a Petri dish containing 5X Phosphate-Buffered Saline. Animals were dissected individually under a dissecting scope and the midgut was removed. Midgut tissue was placed directly into a microcentrifuge tube containing RNAlater, a stabilizing solution, and stored at 4˚C. Midgut tissues from 12 individuals were collected from each stage.

RNA Isolation

The QIAGEN RNeasyprotocol was followed for isolation of RNA from the midgut tissue. All steps were performed at room temperature. The tissue was removed from the

RNAlater and placed in 600µL of Buffer RLT. Tissue was homogenized by grinding and then by passing the solution five times through a 20 gauge needle. The solution was centrifuged for 3 minutes and the supernatant was transferred to a clean microcentrifuge tube. One volume of 70% ethanol was added to the lysate and mixed by pipetting. The sample was placed in an RNeasy mini column and centrifuged for 15 seconds at maximum speed. The flowthrough was discarded. Buffer RW1 (700µL) was added to the column and it was centrifuged again. This process was repeated twice with Buffer RPE (500µL). The tube was then centrifuged for 2 minutes to dry out the silica-gel membrane. The RNA was eluted by adding 50µL of water to the tube and centrifuging for one minute. This process was repeated using the eluate from the previous step. The eluate was collected, 1µL of RNaseOUT was added and the solution was stored at -20˚C. In order to remove any remaining genomic DNA, RNA samples were later treated with DNAse by returning the sample to the column, incubating for 15 minutes with DNAse and then repeating the RNA isolation procedure.

Synthesis of cDNA

cDNA was synthesized from RNA isolated from the midgut using the Invitrogen protocol. The isolated RNA solution (10µL) was combined with 1µL of Oligo(dT) (500 µg/mL) and 1µL of dNTP mix (10mM each). This solution was heated to 65˚C for 5 minutes, quickly chilled on ice and briefly centrifuged. 5X First-Strand Buffer (4µL) and 0.1M DDT (2µL) were added. The contents of the tube were mixed and incubated at 42˚C for 2 minutes. SuperScript II RT (1µL) was added and mixed by pipetting. The solution was incubated at 42˚C for 50 minutes and then heated to 70˚C for 15 minutes. DNA was stored at -20˚C.

Real Time Reverse-Transcriptase PCR

Primers for the *E(spl)* genes *mα, mβ* and *mγ,* the ecdysone-inducible genes *E75B* and *EcR,* and the *rp49* gene, which was used as a positive control, were designed using Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and ordered from Invitrogen.

mγ:	Right (5'-3'): ACCAAGATGGAGGACGACAG
	Left (5'-3'): CAGCCAGCAGAAAAGGAAGT
mβ:	Right (5'-3'): CACTCCACCACCCTCTGAAT
	Left (5'-3'): CATCGTCTCAACTACCTGCAA
mα:	Right (5'-3'): GTTCGAGATTGTCGAGGAGC
	Left (5'-3'): CCAGCTACAGCATCAAGCAG
E75B:	Right (5'-3'): TTTTTGTTGGTGCCAGTGAA
	Left (5'-3'): GTTCAGCTCGTGTTGCTCTG
EcR:	Right (5'-3'): TCTGCACTTCGACAAACGAA
	Left (5'-3'): ACACACGCCTACACATCCAG
rp49:	Right (5'-3'): TAGTGCACTGACCCACTGGA
	Left (5'-3'): GCCAAATGGCTTCTGTTTGT

Real Time PCR was run in 96 well plates on an iCycler running Real Time Detection System Software, version 3.1 (BioRad). Dilutions of genomic DNA, isolated from *Drosophila melanogaster*, were used as the standards. The dilutions of genomic DNA were 0.015, 0.01, 0.005, 0.0025, 0.00125 0.000625 and 0.0003125 µg/µl. Separate sets of standards were prepared for each different gene. Each reaction contained 10 µL of iQ SYBR green supermix, 8 µL water, 1 µL primer mix (10 µM each) and 1 µL DNA sample. The cDNA samples were diluted to a concentration of 0.922 µg/µl. Forty cycles of PCR was performed with a denaturation temperature of 95°C, an annealing temperature of 60°C and an elongation temperature of 72°C. Four replicates were run for each sample.

The log of the concentrations of the standard genomic DNA was graphed as a function of threshold cycle, the cycle at which fluorescent intensity increases above background fluorescence, and the standard curve equation was obtained. This equation was then used to calculate the initial starting quantities of DNA for each gene present in the 5 cDNA samples,

based on the observed threshold cycle. rp49 was run as a positive control for each of the other genes being tested. Based on the fact that rp49 expression remains constant throughout development (Dubrovsky, 2004), the levels of rp49 expression in the different cDNAs were used to normalize the concentration of the other genes.

Results

Bioinformatics Analysis of the *E(spl)* Regulatory Regions

A total of 31 different transcription factor binding sites were identified in the 10 *Drosophila melanogaster E(spl)* genes that were examined here (Table 1). Ttk69 sites were identified in the upstream regulatory region of 3 of the *melanogaster E(spl)* genes; 3 sites in *mγ*, 2 sites in *mα* and 1 site in *m5*. Binding sites for transcription factors that are known to be active during metamorphosis were identified in many of the *E(spl)* genes. These identified transcription factors include those produced by *Hairy, Dorsal, Fushi Tarazu* and *Broad-Complex* genes. *Hairy* binding sites were identified in 9 out of the 10 *E(spl)* genes in *D. melanogaster*, with only *mδ* lacking a binding site. Binding sites for transcription factors from the *Dorsal* gene were located in 8 of the *E(spl)* genes and *Fushi Tarazu* binding sites were identified in 7 *E(spl)* genes. *Broad-complex Z1* and *Z2* binding sites were both found in the upstream regulatory region of 5 and *Z4* was found in 1 of the *melanogaster E(spl)* genes. E74A ecdysone binding sites were identified in the upstream regulatory sites of 4 *melanogaster E(spl)* genes: 1 site each in m2, m3 and m7 and 3 sites in *mγ*.

Many of the binding sites identified in *Drosophila melanogaster* were found to be conserved, or partially conserved, in other species of *Drosophila*. A total of 59 Su(H) sites were identified in the upstream regulatory regions of 10 *melanogaster E(spl)* genes (Table 2). Of these sites, 17 were fully conserved in all of the 9 other *Drosophila* species that were examined. *D. simulans* and *D. yakuba* showed the highest level of conservation, with 45 and 49 fully conserved sites and 6 and 8 partially conserved sites, respectively. *D. grimshawi* had the least conservation and showed only 26 full and 4 partially conserved sites. Similar patterns of site conservation were observed for the N boxes (Table 3) and E boxes (Table 4). Some genes show

site conservation patterns that differ from the general trend. *mγ* and *mα* have high levels of conservation of Su(H) sites; in both genes, 4 out of the 6 sites are conserved in all 9 species. While *m2* also shows a high level of conservation of Su(H) sites, 7 out of the 10 sites are lost in *D. simulans*, and 5 of these sites remain conserved in all of the other species. For most of the 9 other *Drosophila* species, each *E(spl)* gene retains at least one conserved Su(H) site, E box and N box. With only one exception, at least 1 Su(H) site in each gene is conserved for each species. The *m1* gene in *D.grimshawi* is the only gene without a conserved Su(H) site. The N boxes and E boxes are not as well conserved; fully conserved N boxes are not found for *m1* in *.D. mojavensis*, *m3* in *D. virilis* or *D. mojavensis,* or *mα* inD. v*virilis*, *D. mojavensis* or *D. grimshawi* and fully conserved E boxes are not found for m7 in *D. ananassae, D. persimilis* or *D. mojavensis* and *mβ* contains a fully conserved E box only in *D. simulans*.

Conservation across species is not as high for the ecdysone binding sites (Table 5). *m3* and *m7* each contain one ecdysone binding site which is very poorly conserved; the *m3* site is not conserved in any other species and is partially conserved only in *D. simulans* and *m7* has complete conservation only in *D.simulans* and partial conservation in *D. yakuba* and *D. erecta*. The ecdysone binding site in the m2 gene is fully conserved in 5 out of the 9 species and partially conserved in an additional 3 species. The first *mγ* site is also poorly conserved, however, the second and third show full or partial conservation in all species and 6 out of the 9 species retain at least one ecdysone binding site upstream of the *mγ* gene. Bioinformatics results show that while some sites are lost, many are highly conserved across species and most genes retain at least one functional form of each binding site.

Real-Time RT-PCR Detection of Gene Expression During *Drosophila* Gut Development

Real-Time Reverse-Transcriptase Polymerase Chain Reaction (Real Time RT-PCR) was used to quantify the amount of three different *E(spl)* genes during phases of *Drosophila* metamorphosis. The rp49 gene was used as a positive control.

Low concentrations of DNA for *mα* were seen during the wandering third instar larval stage, however there was a 53% increase in concentration from the early to late third instar stage. At puparium formation *mα* concentration increased by 226% and at 2-3 hours post puparium formation it increased by an additional 131% (Figure 6). The concentration of *mβ* was relatively high during the early third instar stage and then dropped by 81% by the late third instar larval stage. The concentration then increased by 311% at puparium formation and increased again by an additional 98% at 2-3 hours post puparium formation to give a final concentration that was 51% greater than the concentration at the beginning of the third instar larval stage (Figure 6). The overall concentration of *mγ* was much lower than that of *mα* and *mβ*. From the early to late third instar larval stage, *mγ* showed a 27% increase. The concentration decreased by 63% at puparium formation and then increased by 23% at 2-3 hours post puparium formation (Figure 7).

Expression of *EcR* and *E75B* was also examined. *EcR* had a high level of expression during the early third instar larval stage. It then decreased by 48% by the late third instar stage, and decreased by an additional 50% at puparium formation. *EcR* then remained relatively constant, increasing by only 9% at 2-3 hours post pupariation (Figure 6). *E75B* showed an overall low concentration of DNA. Expression appears to have increased from the early to late third instar larval stage. However, standard error for the data from the late third instar stage is such that no definite conclusions can be drawn from those numbers. However, there is a decrease in expression of 43% from the early third instar stage to puparium formation. This is

then followed by an increase in expression during the stage 2-3 hours post puparium formation (Figure 7). These results indicate similar expression patterns of *mγ* and ecdysone-inducible genes during development, which suggests a mechanism of direct ecdysone control for the expression of *mγ*.

Activated Notch Significantly Increases Gene Expression

The luciferease assay was performed on S2 cells that were transfected with the luciferase vector containing the *mγ* promoter region. The insertion of the *mγ* promoter region upstream of the *luciferase* gene puts transcription, and consequently protein production, of this gene under the control of the *mγ* promoter. In this way, the luciferase assay, which measures levels of luminescence, was used to report and quantify activation of the *mγ* promoter. Higher levels of luciferase activity reflect higher levels of transcription from the *mγ* promoter. Basal levels of luminescence were very low in the cells transfected with only the luciferase vector. Luminescence drastically increased in cells transfected with the luciferase vector and activated forms of Notch (NICD). In the presence of activated Notch there was a 30 fold increase in luciferase activity (Figure 8). This increase in luminescence corresponds to increased transcription of the gene under the control of the *mγ* promoter. These results indicate that NICD acts to increase *mγ* transcription and also shows that the vector is functioning properly and can now be used for further experiments.

Discussion

In order to gain an understanding of the interactions of transcription factors which act to regulate transcription, I investigated the expression of the *E(spl)* genes in *Drosophila melanogaster* during development. Three different approaches were used to gain an understanding of *E(spl)* gene regulation. Bioinformatics analysis of the upstream regulatory regions of the *E(spl)* genes allowed for identification of conserved transcription factor binding sites that may play a role in gene regulation. The evolutionary conservation of the Ttk69 binding site lends support to the importance of Ttk69 in the control of *mγ* expression. My preliminary transfection results showed that the newly constructed *mγ*promoter-Luciferase vector was activated by NICD and will be an excellent tool to determine whether Ttk69 represses *mγ*. Bioinformatics examinations also revealed conserved ecdysone binding sites in 4 of the *E(spl)* genes. This suggests the possibility that ecdysone acts, either directly or indirectly, to control the expression of the *E(spl)* genes. To further examine this, I used Real Time RT-PCR, to measure the levels of three *E(spl)* genes at different stages of metamorphosis, in order to characterize expression patterns of these genes during development.

Bioinformatics Analysis of the *E(spl)* Regulatory Regions

The high level of conservation of the Su(H) sites, E boxes and N boxes in the 10 *Drosophila* species suggests that these sites are very important in the regulation of *E(spl)* gene expression. The *m1* gene in *D. grimshawi* is the only gene, out of 10 genes each examined in 10 different species, that does not contain a conserved Su(H) site. This implies that at least one Su(H) site must remain functional in order for the gene to be properly expressed. *m1* may have a mutated version of the binding site that differs from the *D. melanogaste*r sequence, but is

nevertheless still functional. Alternatively, *m1* may not be activated by Notch in *D. grimshawi* and may function differently in this species. All of the Su(H) sites appear to be important, however, examination of the conservation of these sites suggests that the most crucial are the two paired Su(H) sites. Six of the *E(spl)* upstream regulatory regions that were examined are identified as containing paired Su(H) sites. Of these, both Su(H) sites are completely conserved in all species for *mδ, mγ, m7and m8*. In *m3* and *m5* one of the paired sites is completely conserved while the other is lost only in *persimilis* for *m3* and *mojavensis* for *m5*. The conservation of the paired Su(H) sites is much higher than the overall concentration of Su(H) sites. The experiments done by Cave et al (2005) showed that interaction between Notch and the proneural proteins required a specific orientation of paired Su(H) binding sites. The high conservation of these paired sites across *Drosophila* species further supports the importance of this specific orientation of sites in the regulation of the *E(spl)* genes.

Many different transcription factor binding sites were identified in the upstream regulatory regions of the 10 *E(spl)* genes. The bHLH transcription factors that are coded for by the *E(spl*) genes play many different roles in *Drosophila* development. The transcription of these genes is controlled by transcription activators and repressors that bind to the upstream regulatory region of these genes. The identification of these specific transcription factor binding sites allows us to gain an idea of what other molecules are active in regulating *E(spl)* transcription. Many of the transcription factor binding sites that were identified in the upstream regulatory regions of the *E(spl)* genes are known to be active during development. The ability to affect expression of multiple genes involved in development gives these transcription factors a tighter control over the nuances of gene expression during this crucial period. The presence of binding sites for these molecules in the regulatory region of the *E(spl)* genes implies a role for

them in the transcriptional regulation of the *E(spl)* encoded bHLH proteins and further implicates these molecules in the control of developmental processes.

Binding sites for *Hairy*, a known transcriptional repressor, were identified in 9 out of the 10 *E(spl)* genes. The *Drosophila* Hairy protein plays vital roles in developmental processes such as myogenesis, somitogenesis, sex determination, vasculogenesis, mesoderm formation, and neurogenesis (for review, Fisher and Caudy, 1998). Hairy has been shown to be involved in body segmentation and bristle patterning during development (Rushlow et al., 1989). Another study using a chromatin profiling method DamID to search for transcriptional targets of *Hairy* in *Drosophila* identified putative targets which implicate a role for *Hairy* in the regulation of the cell cycle, cell growth and morphogenesis (Bianchi-Frias et al., 2004). The presence of *Hairy* binding sites in the regulatory regions of the *E(spl)* genes implicates *Hairy* as a potential transcriptional regulator of these genes and gives it an additional level of control over *Drosophila* development.

A binding site for the Dorsal protein, which is involved in dorso-ventral patterning during *Drosophila* development, was identified in 8 out of the 10 *E(spl)* genes. The maternal dorsal protein is distributed throughout the cytoplasm of unfertilized eggs, however, directly following fertilization it undergoes a nuclear transport process. In the ventral regions, Dorsal is released from the cytoplasm and enters the nucleus, while in dorsal regions it remains in the cytoplasm, thus establishing a dorsal-ventral regulatory gradient (Rushlow, 1989). The Dorsal transcription factor acts in a concentration-dependent manner to initiate the differentiation of embryonic tissues through regulation of zygotically active target genes (For a review, Rusch and Levine, 1996). The presence of Dorsal binding sites in the *E(spl)* genes identifies them as putative

Dorsal target genes. The ability to regulate transcription of the *E(spl)* genes could give Dorsal an additional pathway through which to affect developmental processes.

A binding site for the gene product Fushi tarazu was identified in 7 out of 10 *E(spl)* genes. *Fushi tarazu (ftz)* is a pair rule gene that is expressed early in embryonic development. Like the *even-skipped* gene, it is expressed in vertical stripes, and the two play complementary roles; *even-skipped* is transcribed in odd numbered embryonic segments, while *fushi tarazu* is transcribed in even numbered ones. *Fushi tarazu* has also been shown to play a role in cell fate during *Drosophila* neurogenesis. Creation of *ftz* mutant embryos showed that the absence of *ftz* expression in the central nervous system resulted in abnormal neuron development (Doe\et al, 1988). The overall role of *fushi tarazu* in *Drosophila* development, and specifically its role in neurogenesis, suggest a connection between this and the *E(spl)* genes. The presence of Ftz binding sites in the regulatory region of the *E(spl)* genes, and consequently the potential ability for it to effect transcription of these genes, could be an additional mechanism by which Ftz regulates *Drosophila* development.

Binding sites for three different isoforms of the *broad-complex (BR-C)* genes were located in the upstream regulatory region of the *E(spl)* genes. The three isoforms (Z1, Z2 and Z4) were located in 6, 5 and 1 of the *E(spl)* regulatory regions, respectively. Of the 10 *E(spl)* genes only one, *m5*, does not contain at least one of the broad-complex binding sites. The Broad Complex gene product acts as a regulator of ecdysone-induced responses in many tissues during *Drosophila* metamorphosis. There are four *BR-C* isoforms, which contain a core conserved domain and are alternatively spliced to different zinc-finger DNA binding domains, as the result of post-transcriptional regulation. All of the *BR-C* zinc-finger RNA isoforms are induced by ecdysone (Bayer et al., 1996). By studying *Sgs-4* induction, research has shown that BR-C

directly mediates tissue-specific responses to ecdysone during *Drosophila* metamorphosis (von Kalm et al, 1994). The presence of BR-C binding sites, in addition to the identified ecdysone responsive gene, E74A, binding sites, could give ecdysone an even higher level of control over the expression of *E(spl)*.

The identification of binding sites for the *E74A* ecdysone-inducible gene in the upstream regulatory region of 4 of the *E(spl)* genes suggested that ecdysone may play a role in regulation of these genes. This site was found to be poorly conserved in *m3* and *m7,* but well conserved in *m2* and *mγ*. This could indicate that ecysone directly regulates *m2* and *mγ*, while it may indirectly affect expression of other *E(spl)* genes, such as *mα* and *mβ*. Many of the other transcription factors for which binding sites were identified are known targets of ecdysone regulation. As described above, *BR-C* is also an early ecdysone-inducible gene. Studies have shown that *BR-C* and *E74* play both redundant and synergistic roles during metamorphosis to control ecdysone-regulated gene expression (Fletcher and Thummel, 1995). Both the *BR-C* and *E74* genes are directly induced by ecdysone and they are then in turn capable of regulating transcription of the *E(spl)* genes. The presence of binding sites for both of these genes in the regulatory region of the *E(spl)* genes strongly suggests a mechanism for the control of expression of some of the *E(spl)* genes by ecdysone, while other *E(spl)* genes, which lack binding sites for ecdysone-inducible genes, are likely controlled by other transcription factors. Bioinformatics suggested potential transcription factors which may play a role in *E(spl)* gene regulation. Further evidence for this hypothesis was obtained by quantifying gene expression of *E(spl)* genes during development and comparing it to expression of the factors believed to be responsible for *E(spl)* regulation. In order to determine whether the putative ecdysone binding sites that I identified

were responsive during development, Real Time PCR was used to quantitatively examine gene expression at different stages of *Drosophila* metamorphosis.

Real-Time RT-PCR Detection of *E(spl)* Expression During *Drosophila* Development

Real Time Reverse-Transcriptase PCR allowed for the detection and quantification of *E(spl)* and ecdysone-inducible gene expression during *Drosophila* development in the midgut. Previous experiments have also sought to quantify expression these genes. Andres et al. (1993) used Northern Blot analysis to examine expression of ecdysone-inducible genes through development in RNA isolated from the entire fly. Li and White (2003) performed microarray experiments to examine expression of ecdysone-inducible genes and *E(spl) mα* and *mβ* in RNA isolated only from midgut tissue. I sought to confirm the findings of these studies by quantifying gene expression using the method of Real Time PCR. I also examined expression patterns of *mγ* during development, which had not been previously studied.

Real Time PCR of *mα* and *mβ* showed increases in concentration of both genes at puparium formation and again at 2-3 hours post puparium formation. The concentrations of these two genes are also relatively similar to each other. Microarray experiments done by Li and White (2003) showed that *mα* expression is upregulated starting at puparium formation and continuing to five hours post-pupariation and *mβ* expression increases from two to six hours after puparium formation (Li, and White, 2003). My results indicate an increase in *mβ* expression at puparium formation, while the previous microarray studies do not show an increase of this gene until 2 hours post-puparium formation. The general trend however, indicating an increase in expression of *mα* and *mβ* in the early stages of pupation, is supported both by the microarray results and by my Real Time RT-PCR data. The slight differences between my data and those of

previous experiments could be the consequence of differences in fly staging or could result from the fact that Real Time PCR is more sensitive than microarray analysis and is therefore picking up slight changes that were not detectable by other methods.

In addition to *mα* and *mβ*, I also examined expression of *mγ* because of the presence of E74A binding sites in the regulatory region. This *E(spl)* gene showed a different expression pattern than *mα* and *mβ*, and much lower concentrations overall. *mγ* expression decreased at puparium formation and then showed a slight increase 2-3 hours post-puparium formation. The differences in expression patterns suggest that while expression of *mα* and *mβ* may be regulated by the same transcription factors, there is a different mechanism in place for the control of *mγ* expression.

The upstream regulatory regions of both *mα* and *mβ* contain transcription factor binding sites for the Zeste transvection gene product and the GAGA factor encoded by the *trithorax-like* gene. The regulatory region of *mγ*, however, does not contain binding sites for either of these transcription factors. Both Zeste and GAGA factor (GAF) have been shown to act as enhancers of position effect variegation, the process of inactivating a gene by placing it near to the heterochromatin. By decreasing the tendency of DNA sequences to form heterochomatin, these factors act to upregulate the expression of genes (For review, Wilkins and Lis, 1997). The presence of binding sites for Zeste and GAF in *mα* and *mβ* suggests that they could be factors affecting transcription of these genes. Expression of the *zeste* gene has been observed throughout development in *D. melanogaster*. Experiments using a *zeste-lacZ* hybrid gene found that while *zeste* mRNA is at extremely low levels during the early larval stages of development, expression of the gene rises drastically in the late third instar larva and early pupa (Pirrotta et al., 1988). The expression pattern of *zeste* corresponds with the Real Time RT PCR expression

patterns that were observed for *mα* and *mβ*. Both *mα* and *mβ* showed increased expression at puparium formation, which is the same stage at which *zeste* transcription is upregulated. The presence of binding sites for *zeste* and GAGA, in addition to the simultaneous upregulation of *zeste* and *mα* and *mβ*, suggests that the expression of these *E(spl)* genes during development may be controlled by the Zeste and GAGA transcription factors.

The identification of E74A ecdysone-inducible sites in the regulatory region of *mγ* suggests ecdysone as a possible transcriptional regulator of this gene. Previous studies have examined the expression patterns of ecdysone-regulated genes by examining total RNA as well as midgut-isolated RNA. Real Time results show high levels of *EcR* expression during the third instar larval stage and then a continual decrease during transition to the pupal stage. This corresponds with previous experiments, which showed that expression of *EcR*, both in the midgut and in the whole fly, is high during the third instar larval stage, and then decreases both at puparium formation and in the hours following puparium formation (Li and White, 2003; and Andres, 1993). Real Time results show overall low levels of *E75B* expression. A decrease in *E75B* expression was observed at puparium formation, which corresponds with previous results from the studies done by both Li and White (2003), and Andres (1993). These previous studies show differing expression patterns of *E75B* after puparium formation in the midgut and whole fly. In the midgut, expression appears to drop 2 hours post puparium formation and then increase 3 hours post puparium formation (Li and White, 2003). Analysis of total RNA, however, showed an increase at 2 hours post puparium formation and this expression level then remains constant through 4 hours post pupariation (Andres, 1993). My Real-Time RT PCR results indicate an increase in *E75B* expression in the period 2-3 hours post pupariation. Most likely because I isolated because fly guts over the period of an hour, my data do not show the

change between 2 and 3 hours post pupariation that is evident in the microarray analysis, and is showing only the general increase over this time period.

The observed expression patterns of *mγ* correspond with those of *EcR* and *E75B*. All of these genes show a decrease in expression at puparium formation, which could indicate that expression of *mγ* is, at least in part, dependent on the presence of these ecdysone-regulated genes. The specific binding site that I identified in the upstream regulatory region of *mγ* was for ecdysone-inducible E74A. Experiments show that expression of E74A, both in the whole fly and in the midgut, increases at the end of the third instar larval stage, and then decreases at puparium formation (Li and White, 2003; and Andres, 1993). This corresponds with the expression pattern that I observed for *mγ*. The similar patterns of expression between these genes, and the presence of 3 E74A binding sites in the regulatory region of mγ, suggest that E74A could be a direct regulator of *mγ* expression during development.

Future study of *E(spl)* gene expression during development will require not only replication of the experiments described here, but also real-time PCR analysis of the specific genes that I speculate are involved in *E(spl)* regulation. Thus far, no experiments have examined *zeste* expression in the midgut. In order to provide further evidence for a connection between *zeste* and *E(spl*) *mα* and *mβ*, real-time PCR could be used to quantify *zeste* expression during development of the *Drosophila* midgut. Likewise, Real-Time RT PCR analysis could be used to examine expression of *E74A* in the midgut, in order to definitively link it to expression of *mγ*. In addition to further examination of the mechanisms of transcriptional regulation for the three *E(spl)* genes discussed here, Real-Time PCR should be used to study the expression of other *E(spl)* genes. Binding sites for E74A were also identified in *m2*, *m3* and *m7*. Expression patterns could be examined for these genes to determine whether they, like *mγ*, correspond to

expression levels of the other ecdysone genes. If these genes were found to show similar expression patterns, it would provide further evidence that ecdysone acts directly to regulate transcription of some *E(spl)* genes.

A Potential Role for Tramtrack as a Repressor of *E(spl) mγ*

In order to specifically examine the effect of Ttk on *mγ* expression, I constructed reporter vectors to be used in experiments performed on *Drosophila* S2 cells in a cell culture system. *Drosophila* S2 cells were originally derived from a culture of *Drosophila melanogaster* embryos, and have been previously used as a system to study regulation of *mγ* (Eastman et al., 1997).

We first attempted to insert a 280bp region of the *mγ* upstream regulatory region that contained a pair of Su(H) sites, an E box and a Ttk69 binding site. This region is identical to the wildtype sequence of the upstream regulatory region of *Drosophila melanogaster mγ* (Figure 3) has been shown to regulate endogenous expression of *mγ*. Ligation of the 280 bp *mγ* sequence into the pGL3 vector by sticky-end ligation proved to be unsuccessful. Although digestion of the vector revealed the presence of the *mγ* insert, sequencing results failed to show the presence of the *mγ* sequence in the multiple-cloning region. We hypothesize that the *mγ* sequence was inserting into the vector, but at the incorrect site. It is possible that the repetitive sequences in the pGL3 vector caused recombination between the insert and the vector. After multiple attempts at inserting by sticky-end ligation, a second approach was attempted using the entire 1.2KB upstream regulatory region of *mγ*. A blunt-end ligation of this 1.2Kb sequence into the pGL3 vector was successful and sequencing results confirm the correct placement of the *mγ* regulatory region into the pGL3 vector.

Using the luciferase assay, gene expression was compared in cells containing the luciferase gene under the control of the *mγ* promoter. Levels of luminescence correspond to transcription levels of the gene. Luminescence was greatly increased in cells that contained the NICD. The presence of the Notch intra-cellular domain is crucial for upregulation of *E(spl) mγ* expression. These preliminary results demonstrate that the vector is functioning properly. The next step is to perform transfections with vectors containing Su(H) and Ttk69 sites in addition to NICD. These experiments will specifically examine the effect of Ttk69 on *mγ* expression. If tramtrack is indeed a repressor of mγ expression, we expect to see a dramatic decrease in luminescence in cells transfected with activated Ttk69.

Conclusion

Results from Bioinformatics and Real Time PCR analysis have provided evidence for the control of *E(spl)* gene expression by specific transcription factors. In addition to identification of ecdysone-inducible genes in the regulatory region of 4 *E(spl)* genes, corresponding expression patterns of these genes further supports the hypothesis that some *E(spl)* genes are under direct control of ecdysone. Precise control of gene expression is accomplished by the interaction of a myriad of enhancers and repressors. In addition to the specific factors that are studied here, bioinformatics examination of the regulatory regions of *E(spl)* genes suggests numerous other elements that may aid in the control of *E(spl)* transcriptional regulation. Examination of Ttk69 activity by transfections of *Drosophila* cells allows for the specific investigation of one of these potential regulators. Following further Real Time PCR experiments to solidify the hypothesis for control of *E(spl)* genes by ecdysone and other transcription factors, such as zeste and GAGA,

the transfection system, used here for Ttk69 can be used to provide concrete evidence for *E(spl)* regulation by these proposed transcription factors.

Results of future experiments will serve to further elucidate the mechanisms in place for the control of *E(spl)* gene expression. By discovering how this expression is controlled, we can hope to better understand the process of transcriptional regulation and consequently increase the knowledge of developmental processes in general. The complex interaction between transcriptional activators and repressors serves to mediate the complex process of RNA, and consequently protein, production. Understanding the mechanisms by which this occurs is vital to enhancing knowledge of basic developmental processes. The Notch pathway is known to be a key player in many developmental events and further understanding of the steps in this pathway, and the ways in which it acts to control gene expression, will serve not only to provide insight into the overall process of development, but will also help to understand how abnormalities in this system can lead to a wide range of human diseases.

References

Andres, A.J., and Thummel, C.S. (1994). Methods for Quantitative Analysis of Transcription in Larvae and Prepupae. In *Drosophila melanogaster*: Practical Uses in Cell and Molecular Biology, Goldstein, L. S. B., and Fyrberg, E. A. eds., (San Diego: Academic Press) pp. 565--573.

Andres, A.J., Fletcher, J.C., Karim, F.D., and Thummel, C.S. (1993). Molecular analysis of the initiation of insect metamorphosis: a comparative study of Drosophila ecdysteroid-regulated transcription. Dev. Biol. *160,* 388-404.

Badenhorst, P., Finch, J.T., and Travers, A.A. (2002). Tramtrack co-operates to prevent inappropriate neural development in Drosophila. Mech. Dev. *117,* 87-101.

Balint, K., Xiao, M., Pinnix, C.C., Soma, A., Veres, I., Juhasz, I., Brown, E.J., Capobianco, A.J., Herlyn, M., and Liu, Z.J. (2005). Activation of Notch1 signaling is required for beta-catenin-mediated human primary melanoma progression. J. Clin. Invest. *115,* 3166-3176.

Bayer, C.A., Holley, B., and Fristrom, J.W. (1996). A switch in broad-complex zinc-finger isoform expression is regulated posttranscriptionally during the metamorphosis of Drosophila imaginal discs. Dev. Biol. *177,* 1-14.

Beckstead, R.B., Lam, G., and Thummel, C.S. (2005). The genomic response to 20-hydroxyecdysone at the onset of Drosophila metamorphosis. Genome Biol. *6,* R99.

Bianchi-Frias, D., Orian, A., Delrow, J.J., Vazquez, J., Rosales-Nieves, A.E., and Parkhurst, S.M. (2004). Hairy transcriptional repression targets and cofactor recruitment in Drosophila. PLoS Biol. *2,* E178.

Buescher, M., Yeo, S.L., Udolph, G., Zavortink, M., Yang, X., Tear, G., and Chia, W. (1998). Binary sibling neuronal cell fate decisions in the Drosophila embryonic central nervous system are nonstochastic and require inscuteable-mediated asymmetry of ganglion mother cells. Genes Dev. *12,* 1858-1870.

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M., and Werner, T. (2005). MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics *21,* 2933-2942.

Castro, B., Barolo, S., Bailey, A.M., and Posakony, J.W. (2005). Lateral inhibition in proneural clusters: cis-regulatory logic and default repression by Suppressor of Hairless. Development *132,* 3333-3344.

Cave, J.W., Loh, F., Surpris, J.W., Xia, L., and Caudy, M.A. (2005). A DNA transcription code for cell-specific gene activation by notch signaling. Curr. Biol. *15,* 94-104.

Cooper, M.T., Tyler, D.M., Furriols, M., Chalkiadaki, A., Delidakis, C., and Bray, S. (2000). Spatially restricted factors cooperate with notch in the regulation of Enhancer of split genes. Dev. Biol. *221,* 390-403.

Culi, J., and Modolell, J. (1998). Proneural gene self-stimulation in neural precursors: an essential mechanism for sense organ development that is regulated by Notch signaling. Genes Dev. *12,* 2036-2047.

Delidakis, C., and Artavanis-Tsakonas, S. (1992). The Enhancer of split [E(spl)] locus of Drosophila encodes seven independent helix-loop-helix proteins. Proc. Natl. Acad. Sci. U. S. A. *89,* 8731-8735.

Doe, C.Q., Hiromi, Y., Gehring, W.J., and Goodman, C.S. (1988). Expression and function of the segmentation gene fushi tarazu during Drosophila neurogenesis. Science *239,* 170-175.

Eastman, D.S., Slee, R., Skoufos, E., Bangalore, L., Bray, S., and Delidakis, C. (1997). Synergy between suppressor of Hairless and Notch in regulation of *Enhancer of split* mγ and mδ expression. Mol. Cell. Biol. *17,* 5620-5628.

Fisher, A., and Caudy, M. (1998). The function of hairy-related bHLH repressor proteins in cell fate decisions. Bioessays *20,* 298-306.

Fletcher, J.C., and Thummel, C.S. (1995). The ecdysone-inducible Broad-complex and E74 early genes interact to regulate target gene transcription and Drosophila metamorphosis. Genetics *141,* 1025-1035.

Fletcher, J.C., Burtis, K.C., Hogness, D.S., and Thummel, C.S. (1995). The Drosophila E74 gene is required for metamorphosis and plays a role in the polytene chromosome puffing response to ecdysone. Development *121,* 1455-1465.

Fusse, B., and Hoch, M. (2002). Notch signaling controls cell fate specification along the dorsoventral axis of the Drosophila gut. Curr. Biol. *12,* 171-179.

Go, M.J., Eastman, D.S., and Artavanis-Tsakonas, S. (1998). Cell proliferation control by Notch signaling in Drosophila development. Development *125,* 2031-2040.

Hartenstein, A.Y., Rugendorff, A., Tepass, U., and Hartenstein, V. (1992). The function of the neurogenic genes during epithelial development in the Drosophila embryo. Development *116,* 1203-1220.

Hartenstein, V., and Posakony, J.W. (1990). A dual function of the Notch gene in Drosophila sensillum development. Dev. Biol. *142,* 13-30.

Hitoshi, S., Alexson, T., Tropepe, V., Donoviel, D., Elia, A.J., Nye, J.S., Conlon, R.A., Mak, T.W., Bernstein, A., and van der Kooy, D. (2002). Notch pathway molecules are essential for the maintenance, but not the generation, of mammalian neural stem cells. Genes Dev. *16,* 846-858.

Hitoshi, S., Seaberg, R.M., Koscik, C., Alexson, T., Kusunoki, S., Kanazawa, I., Tsuji, S., and van der Kooy, D. (2004). Primitive neural stem cells from the mammalian epiblast differentiate to definitive neural stem cells under the control of Notch signaling. Genes Dev. *18,* 1806-1811.

Jennings, B., Preiss, A., Delidakis, C., and Bray, S. (1994). The Notch signalling pathway is required for Enhancer of split bHLH protein expression during neurogenesis in the Drosophila embryo. Development *120,* 3537-3548.

Jiang, C., Baehrecke, E.H., and Thummel, C.S. (1997). Steroid regulated programmed cell death during Drosophila metamorphosis. Development *124,* 4673-4683.

Kent, W.J. (2002). BLAT--the BLAST-like alignment tool. Genome Res. *12,* 656-664.

Lee, C.Y., Cooksey, B.A., and Baehrecke, E.H. (2002). Steroid regulation of midgut cell death during Drosophila development. Dev. Biol. *250,* 101-111.

Lengyel, J.A., and Iwaki, D.D. (2002). It takes guts: the Drosophila hindgut as a model system for organogenesis. Dev. Biol. *243,* 1-19.

Li, T.R., and White, K.P. (2003). Tissue-specific gene expression and ecdysone-regulated genomic networks in Drosophila. Dev. Cell. *5,* 59-72.

Ligoxygakis, P., Bray, S.J., Apidianakis, Y., and Delidakis, C. (1999). Ectopic expression of individual E(spl) genes has differential effects on different cell fate decisions and underscores the biphasic requirement for notch activity in wing margin establishment in Drosophila. Development *126,* 2205-2214.

Odenwald, W.F., Rasband, W., Kuzin, A., and Brody, T. (2005). EVOPRINTER, a multigenomic comparative tool for rapid identification of functionally important DNA. Proc. Natl. Acad. Sci. U. S. A. *102,* 14700-14705.

Pagans, S., Pineyro, D., Kosoy, A., Bernues, J., and Azorin, F. (2004). Repression by TTK69 of GAGA-mediated activation occurs in the absence of TTK69 binding to DNA and solely requires the contribution of the POZ/BTB domain of TTK69. J. Biol. Chem. *279,* 9725-9732.

Pirrotta, V., Bickel, S., and Mariani, C. (1988). Developmental expression of the Drosophila zeste gene and localization of zeste protein on polytene chromosomes. Genes Dev. *2,* 1839-1850.

Presente, A., Boyles, R.S., Serway, C.N., de Belle, J.S., and Andres, A.J. (2004). Notch is required for long-term memory in Drosophila. Proc. Natl. Acad. Sci. U. S. A. *101,* 1764-1768.

Rusch, J., and Levine, M. (1996). Threshold responses to the dorsal regulatory gradient and the subdivision of primary tissue territories in the Drosophila embryo. Curr. Opin. Genet. Dev. *6,* 416-423.

Rushlow, C.A., Han, K., Manley, J.L., and Levine, M. (1989). The graded distribution of the dorsal morphogen is initiated by selective nuclear transport in Drosophila. Cell *59,* 1165-1177.

Rushlow, C.A., Hogan, A., Pinchin, S.M., Howe, K.M., Lardelli, M., and Ish-Horowicz, D. (1989). The Drosophila hairy protein acts in both segmentation and bristle patterning and shows homology to N-myc. EMBO J. *8,* 3095-3103.

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). Low-melting-temperature agarose DNA purification. In *Molecular Cloning; A Laboratory Manual, 2nd edition* (ed. N. Ford, C. Nolan and M Ferguson), p.6.30-6.31. Cold Spring Harbor Laboratory Press, Plainview, NY.

Schubiger, M., Wade, A.A., Carney, G.E., Truman, J.W., and Bender, M. (1998). Drosophila EcR-B ecdysone receptor isoforms are required for larval molting and for neuron remodeling during metamorphosis. Development *125,* 2053-2062.

Schweisguth, F. (2004). Notch signaling activity. Curr. Biol. *14,* R129-38.

Schweisguth, F., and Posakony, J.W. (1994). Antagonistic activities of Suppressor of Hairless and Hairless control alternative cell fates in the Drosophila adult epidermis. Development *120,* 1433-1441.

von Kalm, L., Crossgrove, K., Von Seggern, D., Guild, G.M., and Beckendorf, S.K. (1994). The Broad-Complex directly controls a tissue-specific response to the steroid hormone ecdysone at the onset of Drosophila metamorphosis. EMBO J. *13,* 3505-3516.

Wakamatsu, Y., Maynard, T.M., and Weston, J.A. (2000). Fate determination of neural crest cells by NOTCH-mediated lateral inhibition and asymmetrical cell division during gangliogenesis. Development *127,* 2811-2821.

Wang, X.D., Leow, C.C., Zha, J., Tang, Z., Modrusan, Z., Radtke, F., Aguet, M., de Sauvage, F.J., and Gao, W.Q. (2006). Notch signaling is required for normal prostatic epithelial cell proliferation and differentiation. Dev. Biol. *290,* 66-80.

Wei, J., and Hemmings, G.P. (2000). The NOTCH4 locus is associated with susceptibility to schizophrenia. Nat. Genet. *25,* 376-377.

Wilkins, R.C., and Lis, J.T. (1997). Dynamics of potentiation and activation: GAGA factor and its role in heat shock gene regulation. Nucleic Acids Res. *25,* 3963-3968.

The Notch Pathway

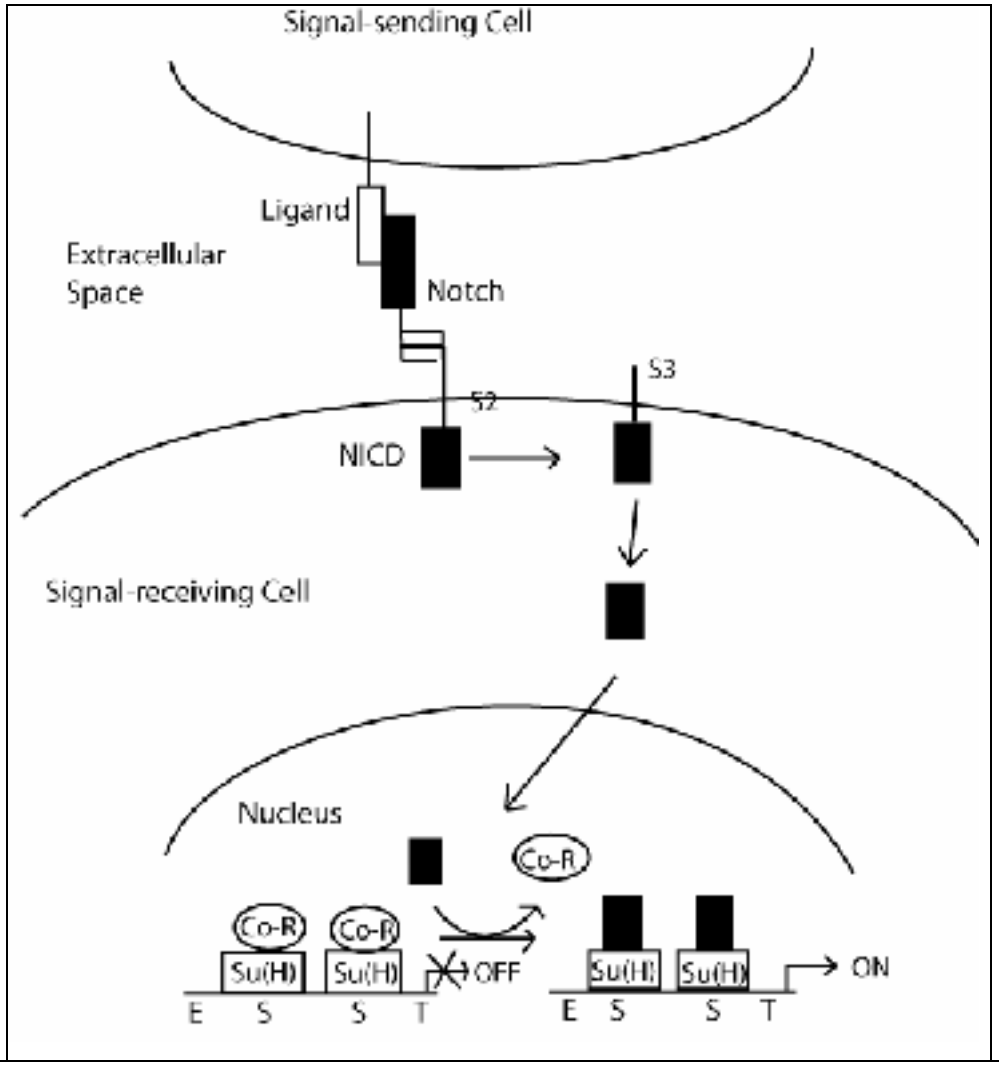

Figure 1. **The Notch Pathway**. Ligand binding with Notch induces proteolytic cleavage, first at the S2 and then at the S3 sites, to release NICD, which is then able to translocate into the nucleus of the receiving cell. In the Nucleus it displaces transcription repressors and associates with transcription factors to upregulate the transcription of target genes.

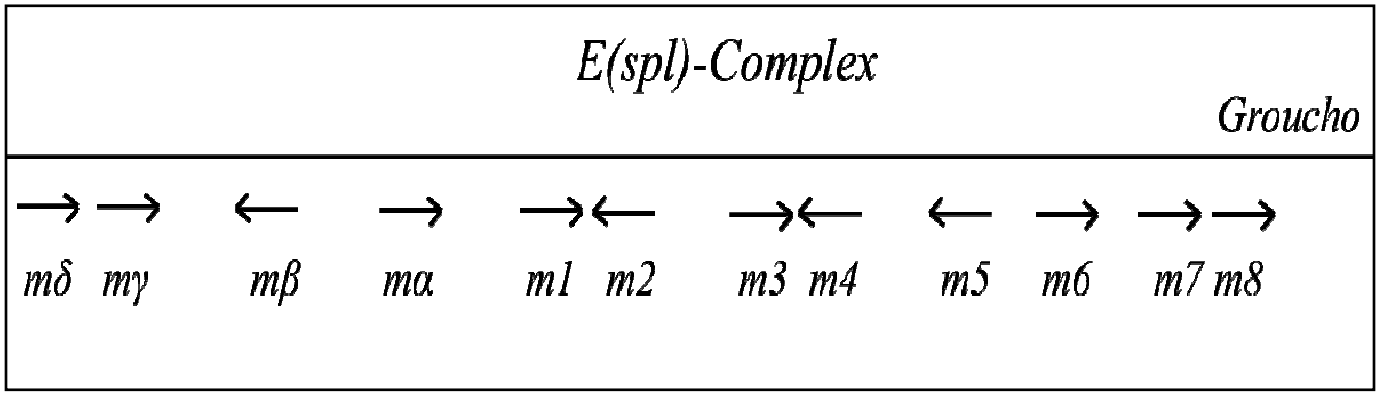

Figure 2. Organization of *E(spl)* gene complex in *D. melanogaster*. Relative position of each of the 12 *E(spl)* genes and *Groucho* is shown. Arrows indicate direction of transcription.

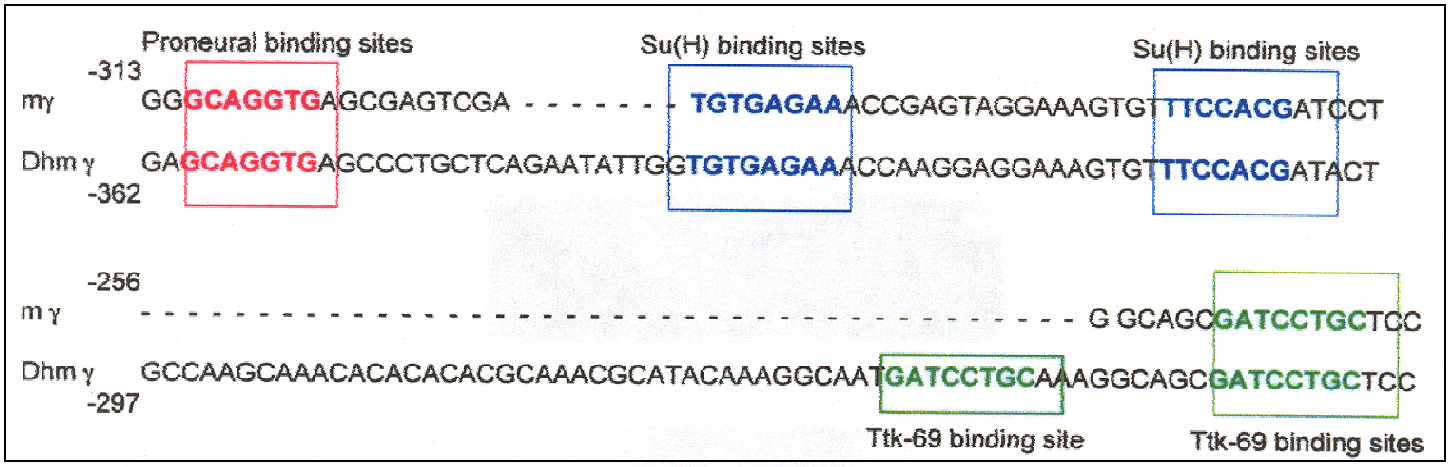

Figure 3. Identification of Ttk-69 binding sites near paired Su(H) sites in the upstream regulatory region of D. melanogaster mgamma (mγ) *D. hydei mgamma* (Dhm γ). Proneural binding sites, Su(H) binding sites and Ttk-69 binding sites are highlighted in bold and in boxes. (D. Eastman NIH grant proposal)

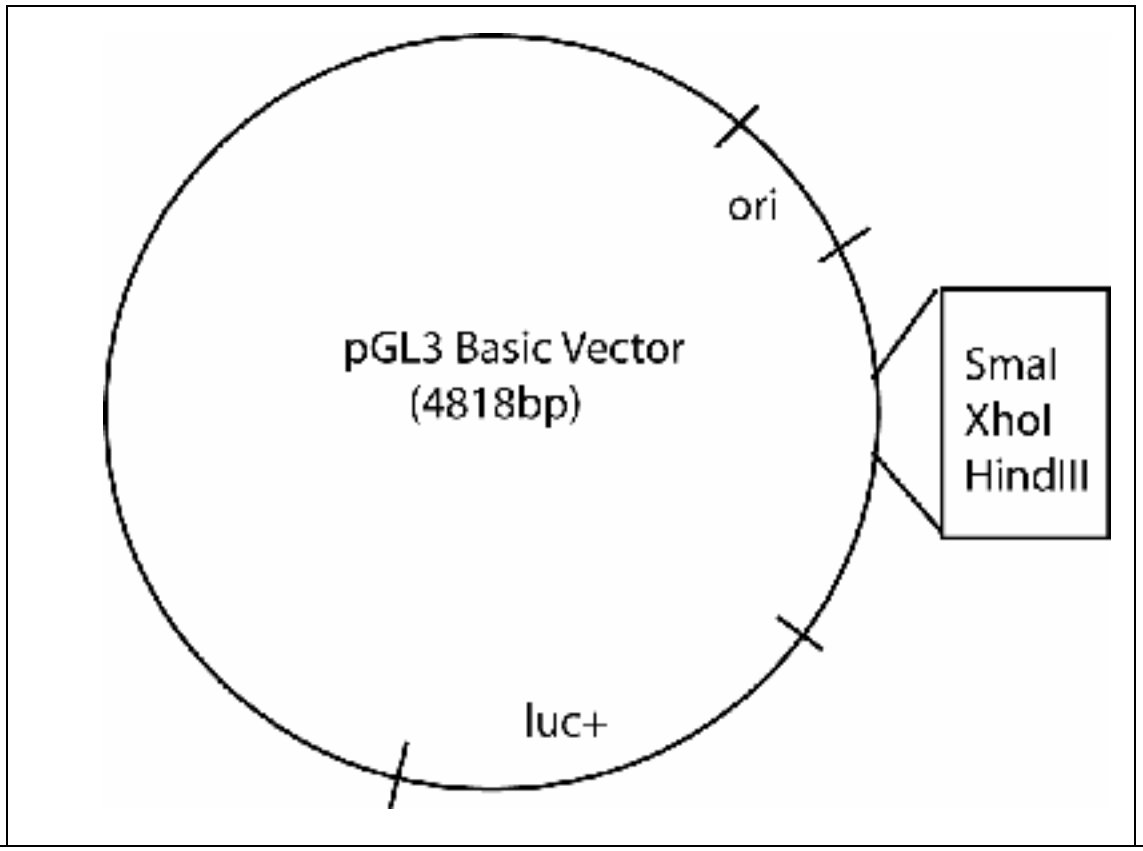

Figure 4. **Map of promega pGL3-Basic vector, showing restriction sites, origin of replication and luciferase gene**

Dissected Midgut Staging of Whole Fly

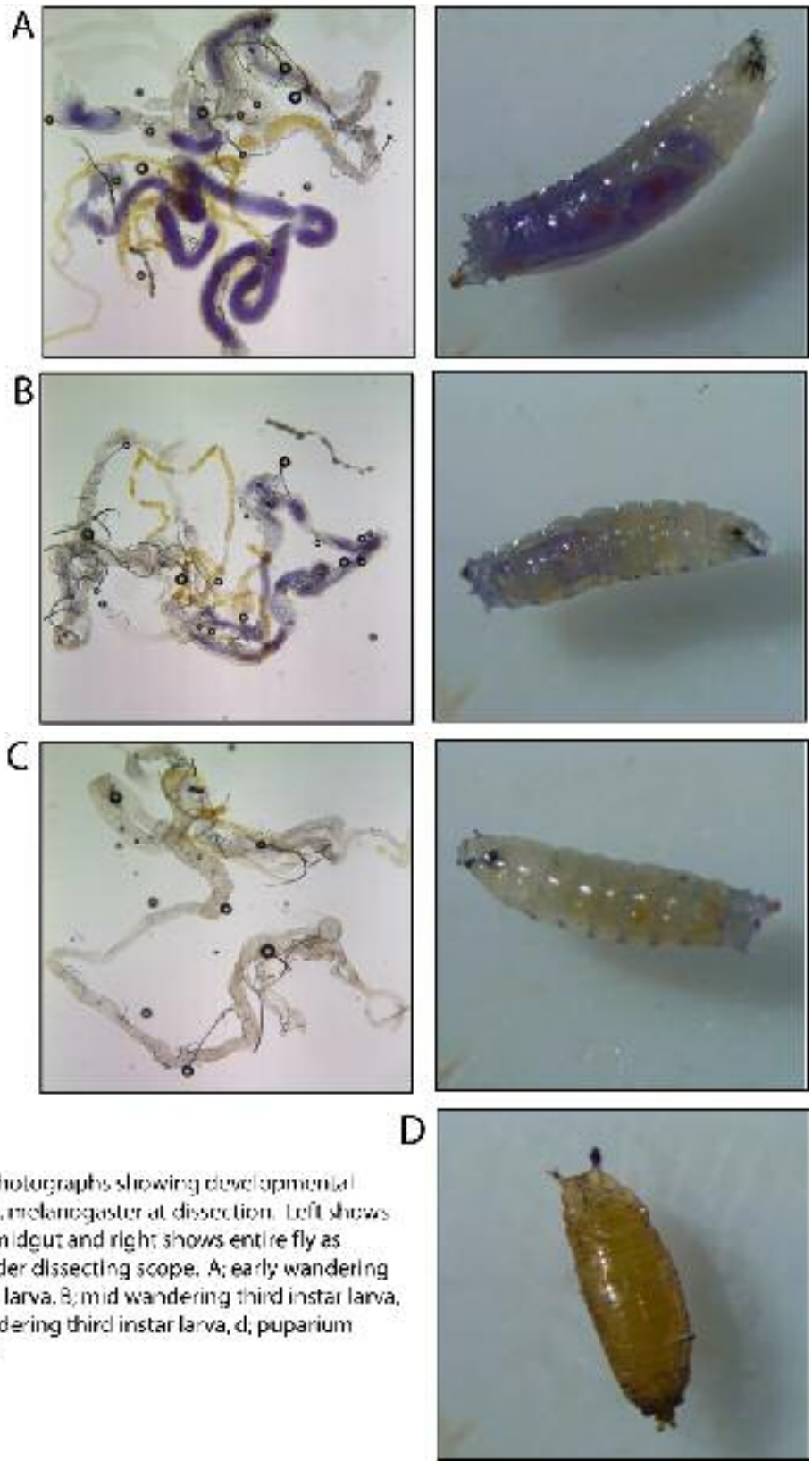

Figure 5. Photographs showing developmental stages of D. melanogaster at dissection. Left shows extracted midgut and right shows entire fly as viewed under dissecting scope. A; early wandering third instar larva, B; mid wandering third instar larva, C; late wandering third instar larva, d; puparium formation.

Table 1. Transcription factor binding sites identified in 10 *Enhancer of Split* genes in *Drosophila melanogaster*. Sites were identified using MatInspector and sequences obtained from UCSC Genome Browser. Number of sites in each gene is indicated as well as the total number of genes that contain each binding site.

	A	B	C	D	E	F	G	H	I	J	K	L
1	**Binding Site**	**m1**	**m2**	**m3**	**m5**	**m7**	**m8**	**malpha**	**mbeta**	**mgamma**	**mdelta**	**# of genes w/ site**
2	tailless (gap gene)	1	1	2		6	1			1		6
3	paired homeodomain	3	3	7	4	4	1	2	5	1	1	10
4	ttk69				1	2		2		3		4
5	hairy, transcriptional repressor	3	2	2	1	3	2	1	2	2		9
6	E74A ecdysone-inducible gene		1	1		1				3		4
7	fushi tarazu	1		4	3	8	6		1	1		7
8	snail zinc finger	1	2	1		3			1	2		6
9	broad-complex Z1	3				2	1		1	1	1	6
10	broad-complex Z2		1	1		1		1			1	5
11	broad-complex Z4					3						1
12	glial cells missing	1			3	2				1		4
13	heat shock factor	6		3	3	9	3	2	3	2	5	9
14	knirp (gap gene)			1		5	2			1		4
15	dorsal, protein for dorso-ventral axis formation		2	1	2	7	3	3		1	1	8
16	Elf-1 (NTF-1), vertebrate homolog CP2				1	1	2		2			4
17	GAGA factor encoded by trithorax-like gene			1		2		1	2			4
18	Adf-1		1	1	8	2			1		2	6
19	Zeste transvection gene product		1	1	1	3	1	2	1			7
20	Hunchback, zygotic zinc finger protein	1		1		2			1			4
21	CF2-II, zinc finger splice variant II	8	18	7	4	3	1				5	7
22	krueppel, zinc finger protein	1	1	1		2	1					5
23	PHO and PHO-like polycomp genes		1								1	2
24	chorion factor 1		1								1	2
25	doublesex, involved in sex determination	1	2	2	1		1					5
26	PAX6 P3 homeodomain binding site	2	2	2		9	1	1				6
27	K50 type homeodomain site		1			1						2
28	crocodile regulator of head development		2	1		3						3
29	deformed, homeotic gene		2	2	2	4	2	2			3	7
30	caudal, homeodomain protein	2	1	4	1	4						5
31	signal transducers and activators of transcription	1		1		2	1	1			2	6
32	T-cell factor, homolog of TCF/LEF	2				4		2				2

Table 2. Conservation of Suppressor of Hairless sites in the *E(spl)* genes in 9 different *Drosophila* species. The species examined were D. *simulans, D. yakuba,D. erecta, D. ananassae, D. persimilis, D. pseudoobscura, D. virilis, D. mojavensis and D. grimshawi* using *D. melanogaster* as the reference.

		simulans	**yakuba**	**erecta**	**ananassae**	**pseudoobscura**	**persimilis**	**virilis**	**mojavensis**	**grimshawi**
M1: 6 sites	# of fully conserved sites	4	3	3	1	1	1	1	1	0
	# of partially conserved sites	2	3	2	1	0	0	0	0	0
M2: 10 sites	# of fully conserved sites	3	9	9	7	8	7	6	5	6
	# of partially conserved sites	7	1	0	0	0	1	1	1	1
Mdelta: 7 sites	# of fully conserved sites	4	3	4	4	3	3	3	3	2
	# of partially conserved sites	2	3	1	0	1	1	1	1	1
M5: 5 sites	# of fully conserved sites	5	5	5	3	3	3	4	3	3
	# of partially conserved sites	0	0	0	2	2	2	0	1	0
M3: 6 sites	# of fully conserved sites	6	5	4	3	3	3	2	2	2
	# of partially conserved sites	0	1	1	1	1	1	1	0	1
Mgamma: 6 sites	# of fully conserved sites	6	6	6	6	5	5	5	5	4
	# of partially conserved sites	0	0	0	0	0	0	0	0	0
M8: 6 sites	# of fully conserved sites	6	6	6	4	4	3	2	3	2
	# of partially conserved sites	0	0	0	2	1	1	0	1	0
M7: 6 sites	# of fully conserved sites	4	5	4	2	3	3	2	2	2
	# of partially conserved sites	2	1	1	0	0	0	0	0	0
Malpha: 6 sites	# of fully conserved sites	6	6	6	4	5	5	4	4	4
	# of partially conserved sites	0	0	0	1	0	0	1	1	1
Mbeta: 1 site	# of fully conserved sites	1	1	1	1	1	1	1	1	1
	# of partially conserved sites	0	0	0	0	0	0	0	0	0
Total: 59 sites	# of fully conserved sites	45	49	48	35	36	34	30	29	26
	# of partially conserved sites	6	8	6	7	5	6	4	5	4

Table 3. Conservation of N Boxes in the *E(spl)* genes in 9 different *Drosophila* species. The species examined were D. *simulans, D. yakuba,D. erecta, D. ananassae, D. persimilis, D. pseudoobscura, D. virilis, D. mojavensis and D. grimshawi* using *D. melanogaster* as the reference.

		simulans	**yakuba**	**erecta**	**ananassae**	**pseudoobscura**	**persimilis**	**virilis**	**mojavensis**	**grimshawi**
M1: 4 sites	# of fully conserved sites	3	3	3	2	2	2	1	0	1
	# of partially conserved sites	1	1	0	1	0	1	1	2	0
M2: 4 sites	# of fully conserved sites	1	3	4	3	2	2	3	1	2
	# of partially conserved sites	0	1	0	0	1	1	0	1	1
Mdelta: 12 sites	# of fully conserved sites	12	12	11	4	8	8	4	5	4
	# of partially conserved sites	0	0	1	2	2	2	2	3	3
M5: 2 sites	# of fully conserved sites	2	2	2	2	2	2	2	2	2
	# of partially conserved sites	0	0	0	0	0	0	0	0	0
M3: 9 sites	# of fully conserved sites	8	6	6	3	2	3	0	0	1
	# of partially conserved sites	0	1	1	1	2	0	3	3	3
Mgamma: 3 sites	# of fully conserved sites	2	1	2	3	2	2	2	2	1
	# of partially conserved sites	1	2	1	0	1	1	1	1	1
M8: 8 sites	# of fully conserved sites	7	6	6	3	2	2	2	2	2
	# of partially conserved sites	0	2	2	3	2	2	2	2	2
M7: 3 sites	# of fully conserved sites	3	2	2	2	2	2	2	2	2
	# of partially conserved sites	0	0	0	0	0	0	0	0	0
Malpha: 2 sites	# of fully conserved sites	2	1	1	1	1	1	0	0	0
	# of partially conserved sites	0	0	1	0	0	0	0	0	0
Mbeta: 3 site	# of fully conserved sites	3	3	3	2	3	3	2	2	2
	# of partially conserved sites	0	0	0	1	0	0	0	0	0
Total: 50 sites	# of fully conserved sites	43	39	40	25	26	27	18	16	17
	# of partially conserved sites	2	7	6	8	8	7	9	12	10

Table 4. Conservation of E Boxes in the *E(spl)* genes in 9 different *Drosophila* species. The species examined were D. *simulans, D. yakuba,D. erecta, D. ananassae, D. persimilis, D. pseudoobscura, D. virilis, D. mojavensis and D. grimshawi* using *D. melanogaster* as the reference.

		simulans	**yakuba**	**erecta**	**ananassae**	**pseudoobscura**	**persimilis**	**virilis**	**mojavensis**	**grimshawi**
M2: 2 sites	# of fully conserved sites	0	2	2	2	2	2	2	2	2
	# of partially conserved sites	0	0	0	0	0	0	0	0	0
Mdelta: 2 sites	# of fully conserved sites	2	2	2	1	1	1	2	2	1
	# of partially conserved sites	0	0	0	1	1	1	0	0	0
M5: 1 site	# of fully conserved sites	1	1	1	1	1	1	1	1	1
	# of partially conserved sites	0	0	0	0	0	0	0	0	0
Mgamma: 2 sites	# of fully conserved sites	2	2	2	2	2	2	2	2	2
	# of partially conserved sites	0	0	0	0	0	0	0	0	0
M8: 3 sites	# of fully conserved sites	3	3	3	2	2	2	2	2	2
	# of partially conserved sites	0	0	0	1	0	0	0	0	0
M7: 2 sites	# of fully conserved sites	2	1	1	0	1	0	1	0	1
	# of partially conserved sites	0	1	1	0	0	0	0	1	0
Malpha: 1 site	# of fully conserved sites	1	1	1	1	1	1	1	1	1
	# of partially conserved sites	0	0	0	0	0	0	0	0	0
Mbeta: 2 sites	# of fully conserved sites	1	0	0	0	0	0	0	0	0
	# of partially conserved sites	1	2	2	1	0	0	0	0	0
Total: 15 sites	# of fully conserved sites	12	12	12	9	10	9	11	10	10
	# of partially conserved sites	1	3	3	3	1	1	0	1	0

Table 5. Conservation of E74A early ecdysone-inducible gene binding sites in the *E(spl)* genes in 9 different *Drosophila* species. The species examined were D. *simulans, D. yakuba,D. erecta, D. ananassae, D. persimilis, D. pseudoobscura, D. virilis, D. mojavensis and D. grimshawi* using *D. melanogaster* as the reference.

		simulans	yakuba	erecta	ananassae	pseudoobscura	persimilis	virilis	mojavensis	grimshawi
M3: 1 site	# of fully conserved sites	0	0	0	0	0	0	0	0	0
	# of partially conserved sites	1	0	0	0	0	0	0	0	0
Mgamma: 3 sites	# of fully conserved sites	2	1	2	1	1	1	0	0	0
	# of partially conserved sites	1	2	1	1	2	2	2	2	2
M7: 1 site	# of fully conserved sites	1	0	0	0	0	0	0	0	0
	# of partially conserved sites	0	1	1	0	1	1	0	0	0
M2: 1 site	# of fully conserved sites	0	1	1	0	0	0	1	1	1
	# of partially conserved sites	0	0	0	1	1	1	0	0	0
Total: 6 sites	# of fully conserved sites	3	2	3	1	1	1	1	1	1
	# of partially conserved sites	2	3	2	2	4	4	2	2	2

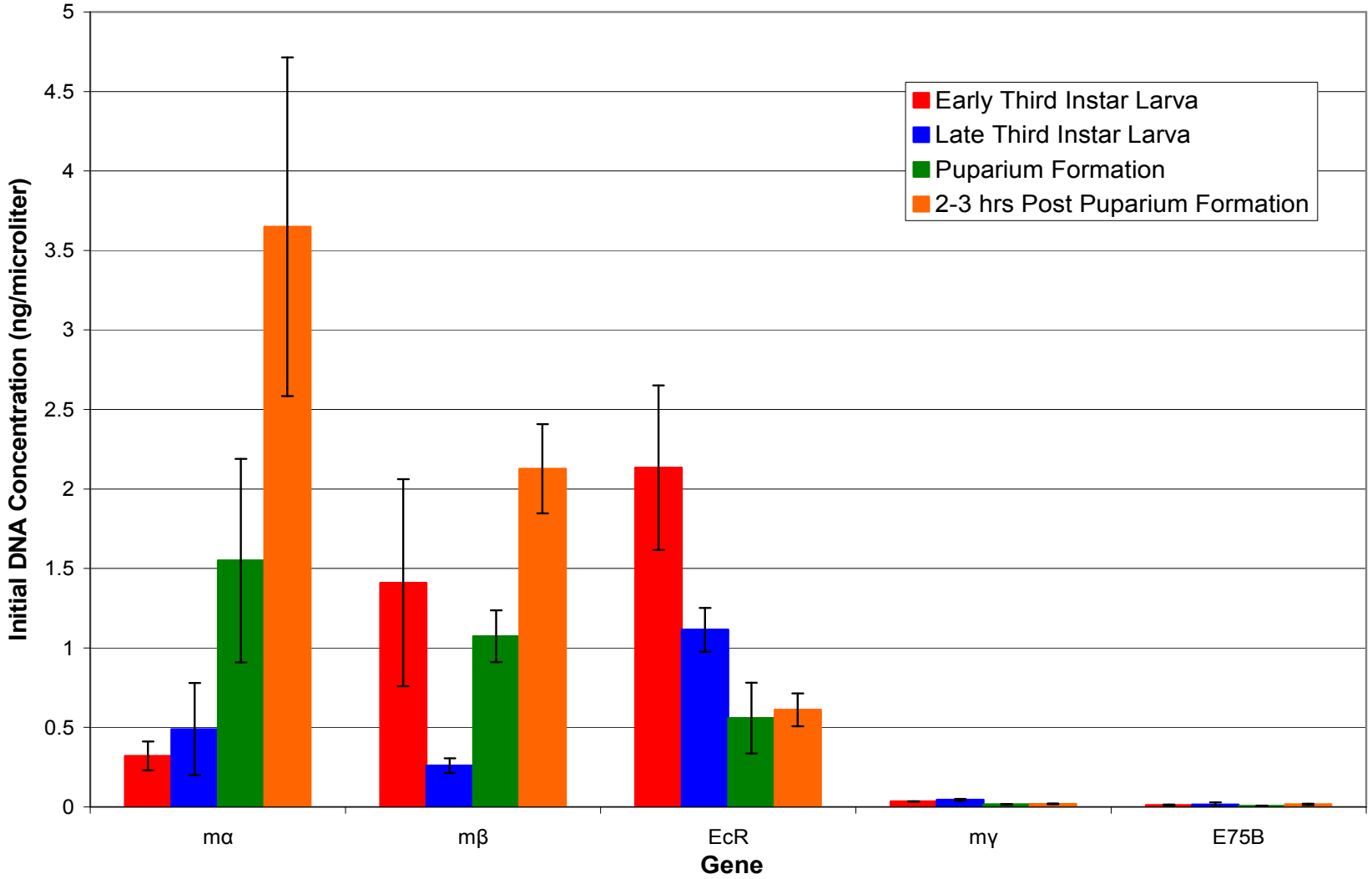

Figure 6. Real-Time RT-PCR data showing concentration of *EcR, E75B, E(spl) mα, mβ* and *mγ* at four different stages of development. RNA isolated from midgut tissue at early third instar stage, late third instar stage, puparium formation and 2-3 hours post puparium formation. Concentrations are adjusted according to rp49. All samples were tested four times and error bars represent the standard deviation of the mean.

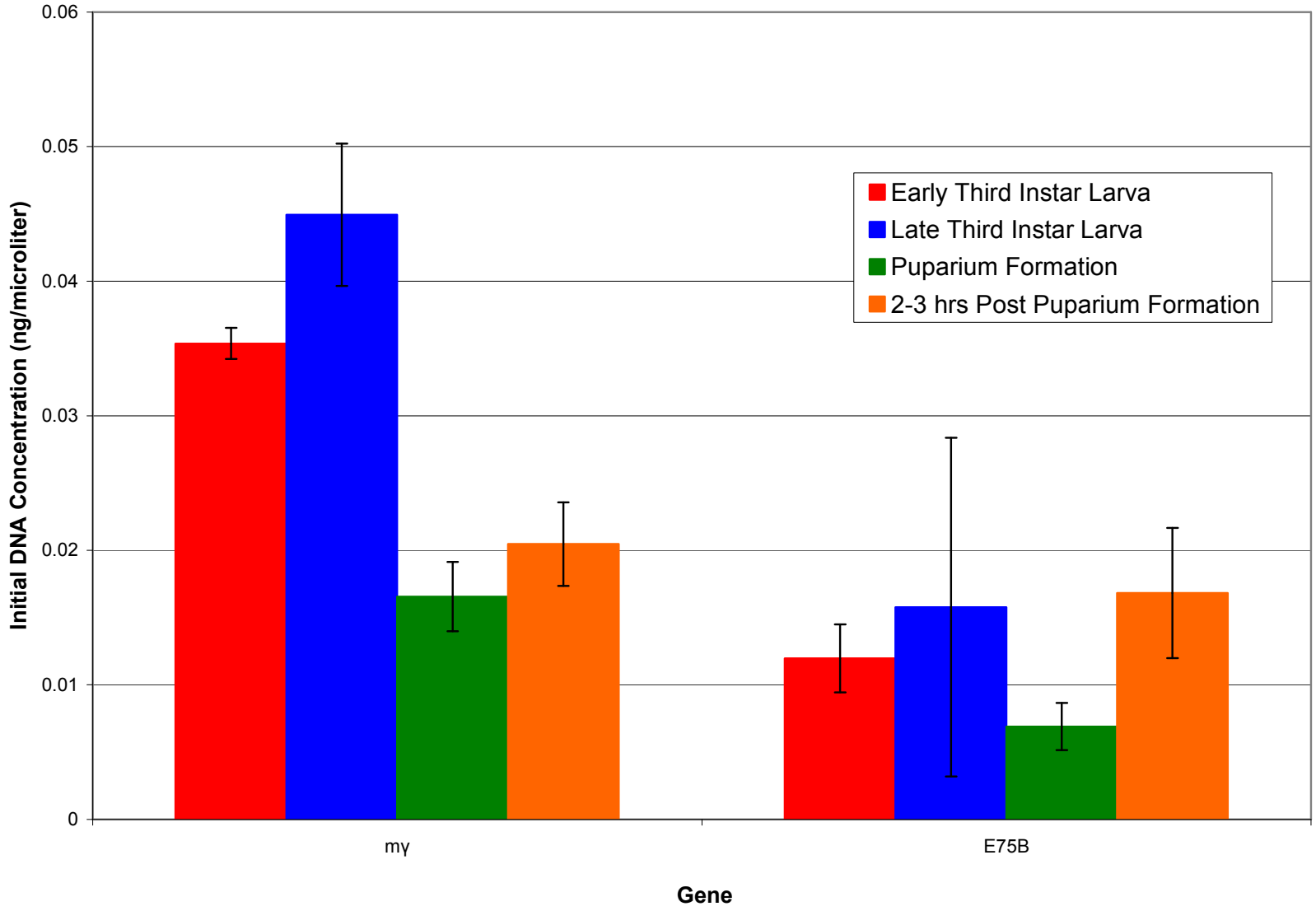

Figure 7. Graph of Real-Time RT-PCR concentration data for *E(spl) mγ* and *E75B*. Data shown is the same as in previous figure, but only the two genes with low concentrations are shown to illustrate the changes in their expression.

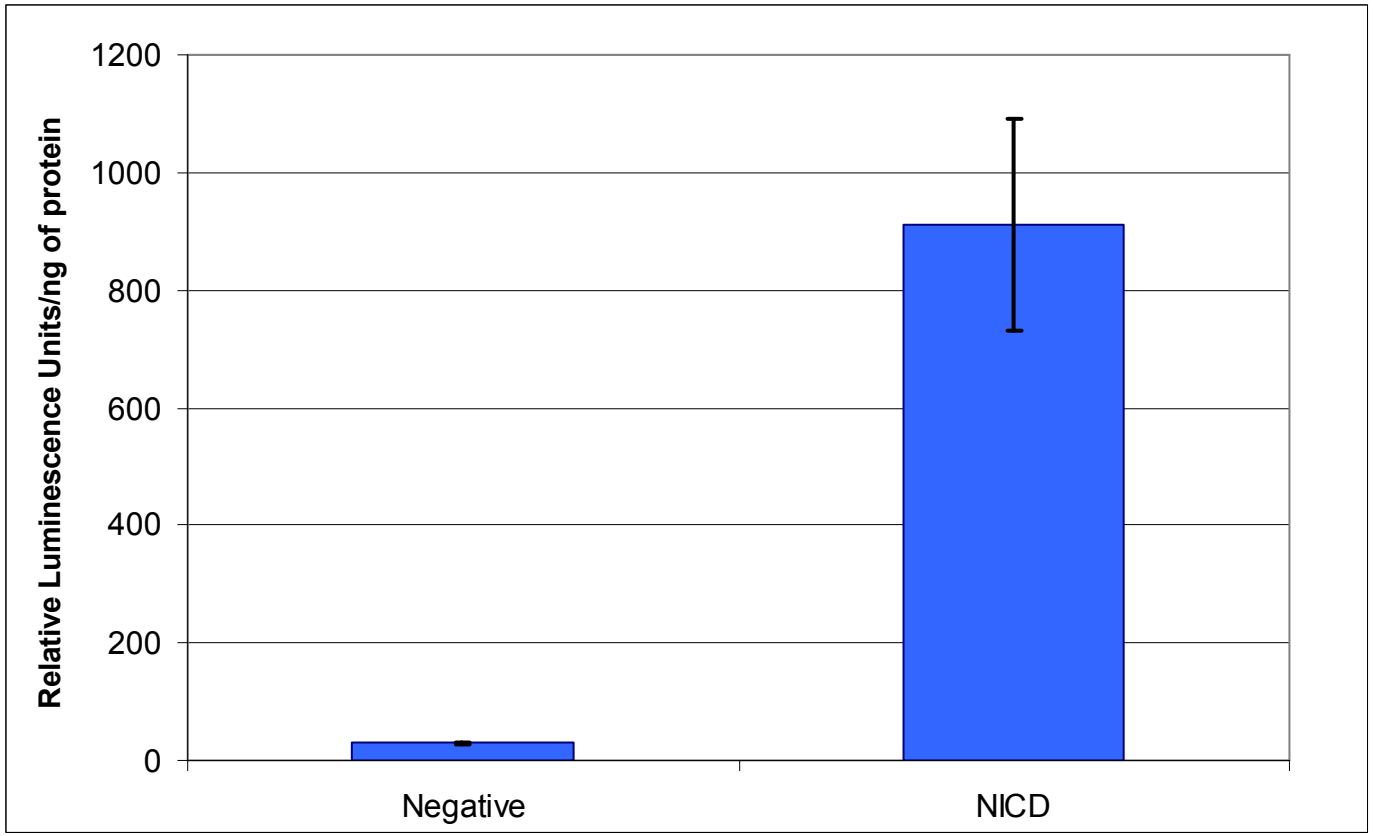

Figure 8. **Luminescence of cells transfected with the luciferase vector containing the mγ promoter region.** "NICD" cells also contain the activated Notch intra-cellular domain. Luminescence was measured on a Veritas Luminometer and Luminescence is graphed as a function of nanograms of protein present in the sample.