2007

Characterization Of Pigment Cell Specific Genes In The Sea Urchin Embryo (strongylocentrotus Purpuratus)

Tricia Stephens
University of Central Florida

STARS Citation

University of
Central Florida

STARS
Showcase of Text, Archives, Research & Scholarship

CHARACTERIZATION OF PIGMENT CELL SPECIFIC GENES IN THE SEA
URCHIN EMBRYO (STRONGYLOCENTROTUS PURPURATUS)

by

TRICIA MARIE STEPHENS
B.S. University of Florida, 2004

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Biology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2007

ABSTRACT

In sea urchin development, cell fate specification appears by the 60-cell stage embryo when several embryonic territories are recognized: the small micromeres, the large micromeres which will generate primary mesenchyme cells, the vegetal2 layer that will give rise to pigment cells, immunocytes, and muscle cells, the vegetal1 layer, as well as the oral and aboral ectoderm. A Delta-Notch signaling event is required for the differential specification of mesodermal cells that will give rise to secondary mesenchyme cells (SMCs).

SMCs produce four cell types: pigment cells, blastocoelar cells, circumesophageal muscle cells, and coelomic pouch cells. Pigment cells are the first to be specified. During primary invagination at the gastrula stage, eight pigment cell progenitors delaminate from the archenteron into the blastocoel. By the pluteus stage, approximately 30 pigment cells are embedded in the ectoderm. Pigment cells produce echinochrome, a napthoquinone pigment.

Previously, several genes in the sea urchin embryo were isolated that are expressed specifically in pigment cell precursors during the blastula stage. The goal of this research was to characterize a subset of these genes, which are highly similar to: the polyketide synthase gene (*Pks*), a sulfotransferase gene (*Sult*), three different members of the flavin-containing monooxygenase gene family (*Fmo*), and the transcription factor glial cells missing (*Gcm*).

Polyketide synthases (PKSs) are a large family of multifunctional proteins mainly found in bacteria, fungi, and plants. They are responsible for the biosynthesis of a variety of polyketide compounds including antibiotics and mycotoxins. In the sea urchin, *SpPks* is required for echinochrome biosynthesis. Flavin-containing monooxygenases (FMOs) are NADPH-dependent flavoproteins mainly found in bacteria, plants, and higher metazoan. They are responsible for

catalyzing the oxidation of several compounds including the detoxification of xenobiotics and activation of numerous metabolites. It is known that *SpFmo1* is required for echinochrome biosynthesis. Sulfotransferases are found from bacteria through higher eukaryotes. These enzymes catalyze the sulfate conjugation of several substrates resulting in either compound detoxification or bioactivation.

I would like to dedicate this thesis to my mother LeAnne and my father Joseph, my mentor Dr.

Calestani, all my friends here at UCF (especially David, Sam, Tamu, and Todd), my beautiful

long-nosed dog Belle, and my fiancé, Erik.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

For multicellular organisms, development begins with a single cell, fertilized egg that gives rise to several different cell types differing in both function and structure. This process of cell specification and differentiation is mainly determined by changes in gene expression. Although the cells of an organism contain the same set of genes, it is the accumulation of different RNA and proteins that is responsible for varying cell types.

Gene expression of eukaryotes is controlled at several different levels from the transcription of RNA to post-translational modifications of a protein (Figure 1). Transcriptional control involves the ability to control when and how frequently a gene is transcribed. This type of control is regulated by the promoter. The promoter of a eukaryotic gene is composed of a basal promoter, which contains the binding sites for RNA Polymerase, and of several cis-regulatory elements containing DNA-binding sites for transcription factors. Cis-regulatory elements are localized at the 5' –end, 3' –end, and within introns. Transcription factors bind to specific DNA sequences and can be repressors or activators of transcription. By DNA looping, repressors and activators interact with proteins bound to the basal promoter resulting in either a repression or enhancement of transcription (Reviewed in Ptashne, 1986).
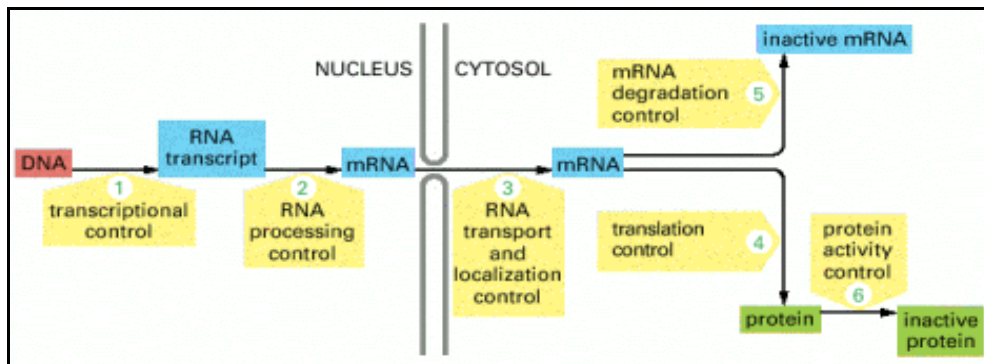
Figure 1: There are six general levels in eukaryotes that control gene expression.

The first level begins at transcription and controlled gene expression continues throughout post-translational modifications of proteins (Alberts et al., 2002).

After transcription, the newly synthesized RNA transcript is not functional and must undergo further processing to become a mature RNA. This is achieved by three processes: 5' capping, 3' cleavage/polyadenylation, and RNA splicing. First, a cap is added to the 5' end of the mRNA transcript that is made of a guanosine nucleotide connected by a 5' to 5' triphosphate linkage. This 5' capping functions as a regulator for export out of the nucleus into the cytoplasm, protects the mRNA molecule from degradation by exonucleases, and promotes the removal of introns. Next, the 3' end is cleaved at a particular cleavage site (AAUAA) and 50-250 adenosines are added at that site. The resulting poly-A tail functions to protect the mRNA transcripts from degradation by the exonucleases, promotes transcription termination, and encourages export of mRNA into the cytoplasm from the nucleus. Finally, RNA splicing is initiated by the spliceosome, a large ribonucleoprotein complex that catalyzes two transesterification reactions responsible for excising introns and joining together exons.

In eukaryotes, there are simple and complex transcription units. Simple transcript units generate a single type of mRNA which encodes a single protein. On the contrary, complex transcription units generate transcripts processed in alternative combinations by the use of alternative poly-A tails or splice sites that result in different mRNAs (alternative splicing). RNA processing control regulates gene expression by determining how the transcript is spliced and processed. One way that alternative splicing is regulated is by proteins binding to sequences next to splice sites, essentially blocking the access to splicing factors. Another way that alternative splicing is regulated is by splicing activators, proteins that interact with splicing factors causing an increase in specificity to the regulated splice site.

Once the mRNA is functional, it can be actively transported to the cytoplasm. Mature mRNA transcripts combine with a variety of proteins in order to form ribonucleoprotein (RNPs) complexes. RNPs that posses a fully processed mRNA are then exported out of the nucleus into the cytoplasm. RNA transport and localization control regulates the timing of when mRNAs are exported, as well as decides where those mRNAs will be localized within the cytoplasm. mRNAs are often directed by proteins that bind to their 3' untranslated region. These proteins combine with motor proteins which allow the mRNAs to be transported to specific locations within the cytoplasm (Reviewed in Kindler et al., 2005).

Once the fully processed mRNA reaches its destination, translational control regulates which mRNAs will be translated into protein by the ribosome. Translational control can be regulated by several different mechanisms. One way this is accomplished is through repressor proteins which bind to specific mRNA sequences in order to block translation. Another way translation is regulated is by controlled polyadenylation. Untranslated mRNAs can be stored with

3

shortened poly-A tails (approximately 20 nucleotides) and when they are needed their poly-A tail is lengthened (to several hundred nucleotides).

Another type of gene expression regulation is mRNA degradation control which involves the ability to selectively destabilize mRNA molecules resulting in degradation. In eukaryotes, there are two main pathways for degradation and sequences within each mRNA molecule decide which pathway. The more common pathway of the two entails the shortening of the poly-A tail by exonucleases in the cytoplasm (Alberts et al., 2002). Once the poly-A tail has been shortened a certain amount (roughly 30 adenosines remaining), the 5' end is decapped which induces instant degradation (Alberts et al., 2002). Furthermore, some mRNAs contain binding sites in their 3' UTR that allow for specific proteins to bind that have the ability to decrease or increase poly-A shortening (Alberts et al., 2002). The second pathway in mRNA degradation involves the direct response of endonucleases. These mRNA contain specific sequences in their 3' UTR that the endonuclease recognizes. Upon recognition, endonucleases cleave the poly-A tail from the mRNA resulting in rapid degradation.

Lastly, protein activity control involves the process of activating/inactivating, degrading, or compartmentalizing protein molecules after translation (Alberts et al., 2002). This type of post-translational control is regulated by several different mechanisms including protein phosphorylation and proteolytic degradation. Protein phosphorylation involves the addition of a phosphate group to an amino acid side chain by kinases. Due to the negative charge of the phosphate group, this process creates a conformational modification of the protein which directly affects the binding of ligands (Alberts et al., 2002). Phosphorylation sites can also be recognized as binding sites by other proteins through their modules, small protein domains that have binding sites for connecting to phosphorylated peptides in other proteins (Alberts et al., 2002).

Dephosphorylation is the removal of the phosphate group by a protein phosphatase and has the ability to restore the protein back to its initial conformation and activity (Alberts et al., 2002).

Proteolytic degradation is a mechanism that identifies and eliminates atypical proteins and it also functions to grant shortened half-lives to particular "normal" proteins when their concentration must be altered immediately. There are three types of enzymes involved in targeting proteins for degradation: ubiquitin-activating (E1), ubiquitin-conjugating (E2) and ubiquitin-ligating (E3) enzymes. Ubiquitin is a small, highly conserved protein that tags target proteins by covalently attaching a lysine residue. A succession of ubiquitin molecules form at the attachment site which is then recognized by proteasomes, large multi-protein complexes that degrade proteins.

This study will focus on the characterization of five pigment cell specific genes in the sea urchin embryo. This research will provide ground work for future studies of gene expression regulation. In the recent years, the sea urchin embryo has become a robust model system for understanding transcriptional regulation. This has been accomplished by large-scale cis-regulatory studies and characterization of gene regulatory networks (GRNs) that underlie sea urchin development. In cis-regulatory regions, transcription factors can bind in a combinational fashion that instructs the basal promoter, thus determining the rate of transcription. Promoter structure is generally studied in two different ways, a blind approach and a more directed approach.

A blind approach typically involves serial deletions of a promoter linked to a reported gene (e.g. GFP). An example in sea urchins is the study of the promoter of cyclophilin 1(*Sp-cyp1*), a gene exclusively expressed in embryo skeletogenic cells (Amore and Davidson, 2006). In order to identify cis-regulatory modules responsible for *Sp-cyp1* expression, they deleted

portions of the *Sp-cyp1* upstream of the first exon (Amore and Davidson, 2006). The excised

promoter was then integrated into GFP expression vectors and injected into eggs. The embryos

were monitored throughout development to determine the regulatory activity of each promoter

fragment and to determine the minimal DNA region required for the normal *Sp-cyp1* expression

(Amore and Davidson, 2006).

A more direct approach for studying promoter structure applies comparative genomics to

identify cis-regulatory modules. This is accomplished by utilizing computational methods that

compare the genomic DNA of one species to the orthologous region of another species (Yuh et

al., 2002). By doing so, it is possible to identify conserved sequences that may or may not

contain cis-regulatory elements. For instance, Yuh (2002) computationally compared the

genomic sequence surrounding the *otx* gene for *S. purpuratus* and *L. variegatus*. In these species,

otx is essential for embryonic endomesoderm specification. Seventeen partially conserved

sequences were amplified from the *S. purpuratus* BAC DNA by PCR, cloned, integrated into a

reporter vector, injected into eggs, and then the regulatory activity was monitored throughout

development (Yuh, 2002). Out of the 17 partially conserved constructs, 11 exhibited spatially

restricted regulatory activity and the remaining six constructs were inactive.

The sea urchin gene regulatory networks (GRN) for endoderm and mesoderm

specification has been extensively characterized (Reviewed in Davidson et al., 2002). The

network has over 40 genes as a result of large-scale perturbation analyses, computational

analyses, genomic data, cis-regulatory analyses and developmental embryology studies.

Furthermore, interactions at each node can be directly confirmed at the genomic sequence by cis-

regulatory analyses. Although the GRN for endo-mesoderm specification has already been

extensively characterized, a considerable amount of research remains for characterizing all of the

cis-regulatory modules for each gene and for determining the direct gene regulatory relationships.

The overall goal of this research is to gain further knowledge on the molecular and cellular mechanisms that regulate cell-type differentiation in the sea urchin, particularly of pigment cells. During development, cells become committed to a particular fate. This process begins with specification, a term used to describe a unique set of gene expression that distinguishes a specified spatial territory. During this time, cells are not permanently defined and continue to migrate prior to differentiation, a term used to describe a cell which acquires a "cell type" due to spatial inductive processes such as transcription regulators of differentiation gene batteries. A gene battery is a set of co-regulated genes which are activated by the same set of cis- and trans-acting regulators.

Currently, very little information is known about the genetic pathways involved in pigment cell specification and differentiation. In order to understand the molecular and cellular mechanisms that are essential for specification and differentiation, a genetic screen was done to isolate genes responsible for these processes. This research characterizes five of the isolated pigment cell specific genes: a polyketide synthase gene (*Pks*), a sulfotransferase gene (*Sult*), and three different members of the flavin-containing monooxygenase gene family (*Fmo*). Collectively, these genes belong to a putative gene battery.

Essentially, the aims of this work are to:

- Annotate each gene (transcript length, number of exons, conserved domains)
- Provide a detailed gene expression profile for each gene
- Determine whether or not there are any alternative spliced isoforms

- Identify putative orthologues of these genes in other model organisms and evaluate their evolutionary history

In order to accomplish these aims, the following techniques were used: computational analyses initially of the BAC clone and then of the fully assembled sea urchin genome, real time quantitative polymerase chain reaction (QPCR), northern analysis, whole mount *in situ* hybridization (WMISH), and phylogenetic analyses. Collectively, this research provides ground work for future studies of gene expression regulation and function of this pigment cell differentiation gene battery. Most importantly, the detailed gene expression profile of these genes will serve as a reference for comparing the effects caused by mutagenesis of the promoter of these pigment cell genes and ultimately characterize the promoter architecture of this pigment cell gene battery.

Sea Urchin Embryo Development

As the sperm enters the egg, the fertilization membrane elevates off the surface of the egg. The first embryonic cleavage is meridional (north – south). The second cleavage is also meridional but at a right angle to the first cleavage plane, thus passing through the animal and vegetal poles. The third cleavage is equatorial (east – west) and perpendicular to the first two cleavage planes, thus splitting the embryo into an animal and vegetal hemisphere. The result of the first three cleavages is an egg divided equally into eight blastomeres (Figure 2).
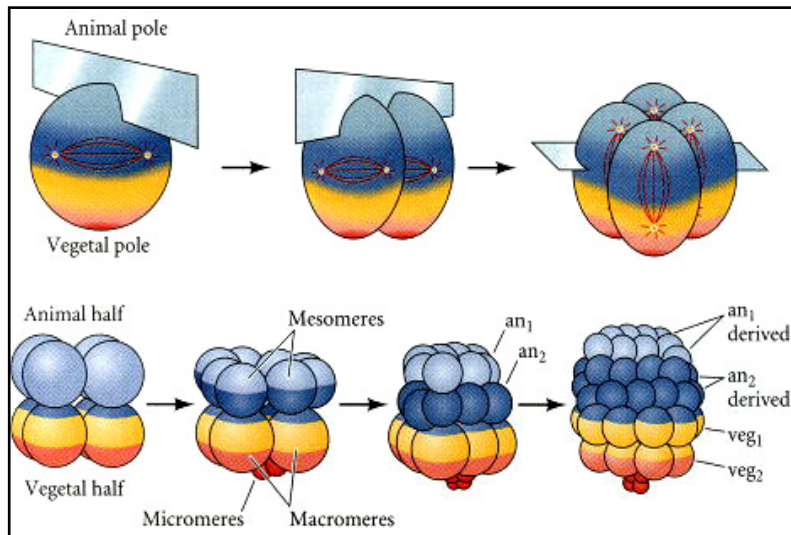
Figure 2: The first six cleavages of sea urchin development.

The first row represents the first three cleavages which result in an embryo with an animal and vegetal hemisphere. The second row represents the fourth through seventh cleavages which result in a 60-cell stage embryo (Gilbert 2000).

At the fourth cleavage, the four cells of the animal tier divide meridionally into eight equally sized blastomeres called mesomeres. The vegetal tier undergoes an unequal equatorial cleavage to produce four large cells called macromeres, as well as producing four smaller sized cells called micromeres (Summers et al., 1993). As the 16-cell embryo undergoes cleavage, the eight mesomeres cleave into two animal tiers, an1 and an2. The macromeres divide meridionally and generate a tier of eight cells directly below an2. The micromeres also divide forming a small cluster beneath the larger tier called the vegetal tiers. Cleavage furrows of the sixth division are equatorial, followed by a meridional cleavage in the seventh division, rendering a 128-cell blastula.

9

By the 60-cell stage embryo, several embryonic territories are already specified: the small micromeres, the large micromeres which will generate primary mesenchyme cells, the veg2 layer that will give rise to pigment cells, immunocytes, and muscle cells, the veg1 layer, as well as the oral and aboral ectoderm (Logan and McClay, 1997) [Figure 3]. The animal half of the embryo (oral and aboral ectoderm) gives rise to the ectoderm and ultimately to the larval epidermis and its' associated neurons. The veg1 layer gives rise to cells that can contribute to the ectoderm or endoderm. The veg2 layer produces secondary mesenchyme cells (SMCs), cells that contribute to the endoderm (precursors of the gut), and cells that contribute to the coelom (precursors of the body wall). The small tier of micromeres will give rise to the primary mesenchyme cells (PMCs) that are responsible for forming the larval skeleton. The large tier of micromeres will eventually gives rise to cells that contribute to the coelomic pouches (Logan and McClay 1997; Logan and McClay 1999).
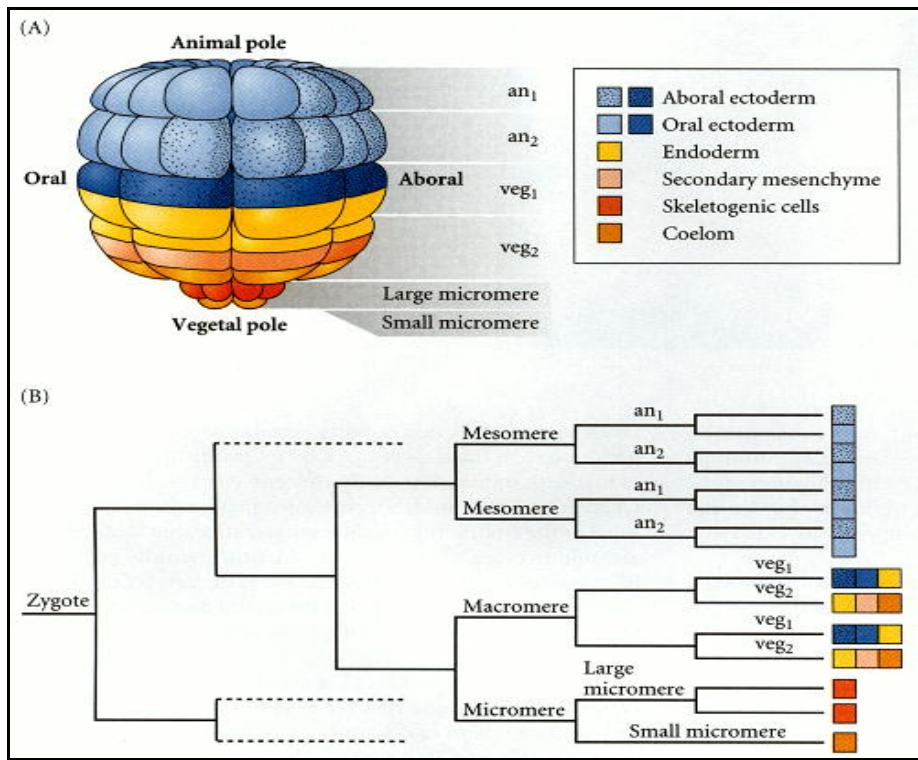
Figure 3: Fate map of the sea urchin, *Stronglyocentrotus purpuratus*.

At the 60-cell stage embryo, several embryonic territories are already recognized: the oral and aboral ectorderm (various colors of blue), the endoderm (yellow), precursors of secondary mesenchyme cells (pink), precursors of primary mesenchyme cells (large micromeres as shown in red), and small micromeres (orange) [Logan and McClay 1999; Wray 1999].

SMC Specification and Differentiation

By late blastula stage, the cells from the veg2 layer begin to thicken and form the vegetal plate, a disc of elongated epithelial cells which will play a role in primary invagination at gastrulation (Reviewed by Davidson et al., 1998). The centermost region of the vegetal plate is identified as SMCs (Ruffins and Ettensohn, 1993; Ruffins and Ettensohn, 1996) [Figure 4].

11

SMCs produce four cell types: pigment cells, blastocoelar cells, circumesophageal muscle cells, and coelomic pouch cells (Cameron et al., 1991; Ruffins and Ettensohn, 1996) [Figure 4]. At mesenchyme blastula, all of the SMC precursors are positioned between 0 and 30° of the vegetal plate, however most of these cell types seem to be differentially dispersed (Ruffins and Ettensohn, 1996). For instance, pigment and muscle cell precursors are distributed almost two times more frequently in the 15-30º region. On the other hand, precursors of blastocoelar cells are distributed 1.4 times more frequently in the 0-15° region. Precursors of coelomic pouch cells are equally distributed.
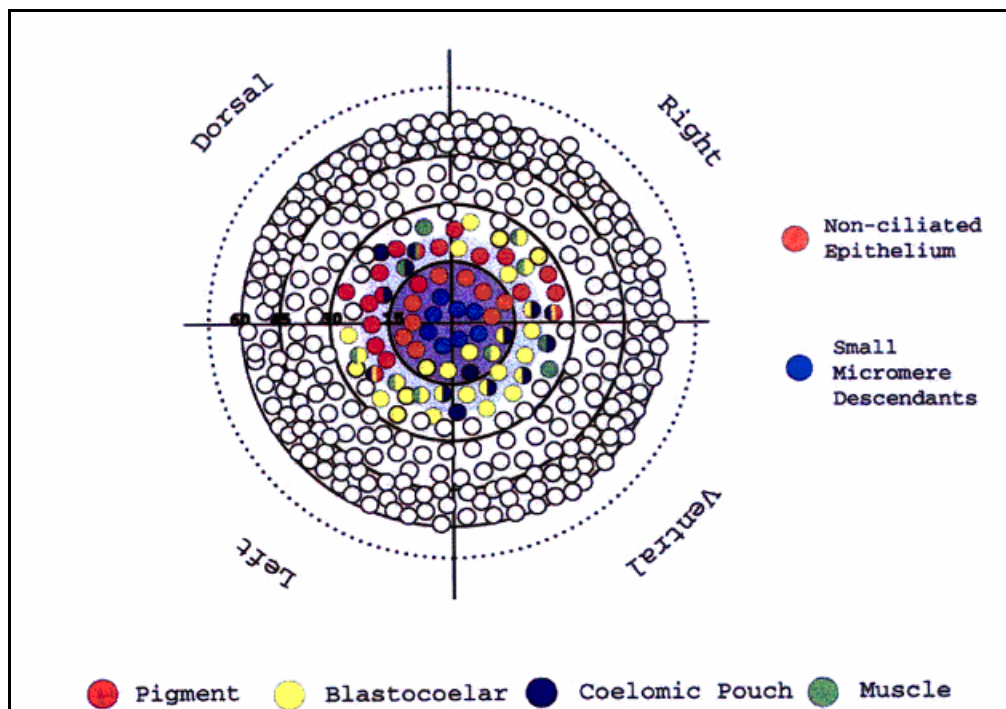
Figure 4: Fate map of SMC precursors at mesenchyme blastula from a vegetal view.

The map begins at the centermost region of the vegetal plate and extends out 60° through the presumptive endoderm and ectoderm (Ruffins and Ettensohn 1996).

Currently, the mechanism behind the specification of the four cell types remains mostly enigmatic. It is known that SMC specification takes place at the blastula stage between the eighth to tenth cleavages (Hörstadius, 1973; Cameron et al., 1991; Ruffins and Ettensohn, 1993; Ruffins and Ettensohn, 1996; Sherwood and McClay, 1999; McClay et al., 2000). A Delta-Notch signaling event is required for the differential specification of endodermal and mesodermal cells which will give rise to SMCs. The Notch signaling pathway is a highly conserved cell-cell communication mechanism. The pathway is involved in several stages of development including cell specification, proliferation, and differentiation. In response to ligand activation (usually by binding to ligands of the Delta or Serrate class), Notch receptors undergo a cascade of proteolytic events which bring about the release of the Notch intracellular domain (NICD). NICD translocates from the plasma membrane to the nucleus where it combines with the C-repeat binding factor 1 (CBF1) repressor complex/ Suppressor of Hairless (SuH) to form an activator complex. The NICD/CBF1/SuH activator complex then up-regulates gene expression (Reviewed in Yoon and Gaiano, 2005).

In sea urchin embryos, the Notch signaling pathway is activated during the blastula stages (Sherwood and McClay, 1999). Delta ligand signals from the large micromeres and is received by the Notch receptor of the neighboring macromeres. Sherwood and McClay (1999) injected eggs with an overexpression of Notch lacking nearly all of the intracellular domain. When the Notch receptor is blocked, SMC formation is drastically reduced creating a diffusion of the endodermal territory into the presumptive SMC territory. There was a 60% decrease in muscle cells, blastocoelar cells, coelomic pouch cells, and almost a 100% depletion of pigment cells.

Pigment cells are the first SMC type to be specified. During primary invagination at the gastrula stage, eight pigment cell progenitors delaminate from the archenteron into the blastocoel

(Reviewed by Davidson et al., 1998). In *S. purpuratus*, during primary invagination, there are approximately 30 pigment cells embedded in the ectoderm (Gustafson and Wolpert, 1963; Gibson and Burke, 1985; Kominami et al., 2001). After the gastrula stage, pigment cells produce echinochrome, a napthoquinone pigment (McLendon, 1912; Kuhn and Wallenfells, 1940; Griffiths, 1965). Two different types of echinochrome are synthesized, type A and B. Echinochrome A is found in the larva as well as the adult spines and tests, while echinochrome B is found only in the larva (Tyler, 1939; Fox and Scheer, 1941; Lederer, 1952).

Unfortunately, the role and function of sea urchin embryonic pigment cells remains unknown. Several hypotheses have been proposed but remain inconclusive. One such hypothesis suggests that pigment cells take part in photoreception. In support of this hypothesis, chromatophores of the sea urchin *Centrostephanus longispinus* respond to light-induced alterations by drastically altering pigment cell shape and granule displacement (Weber and Dambach, 1974; Gras and Weber, 1977; Weber and Gras, 1980). Similar results were also observed in the melanophores of amphibians and fish (Wise, 1969; Schliwa and Bereiter-Hagn, 1973).

Another hypothesis about the possible function of pigment cells is a possible role in immuno-defense. *SpPks* encodes a polyketide synthase (PKS). PKSs are a large family of multifunctional proteins mainly found in bacteria, fungi, and plants. They are responsible for the biosynthesis of a wide variety of polyketide compounds that cover a broad spectrum of biological activity including antibiotic, antitumor, antifungal and immunosuppressive. It is known that *SpPks* is required for the biosynthesis of echinochrome (Calestani et al., 2003). Service and Wardlaw purified echinochrome from coelomic fluid in *Echinus esculentus* and

found that it was bactericidal towards gram-positive and gram-negative marine bacteria (Service and Wardlaw, 1984).

The other three SMC cell types are specified during gastrulation and delaminate from the archenteron tip during secondary invagination. Blastocoelar cells are free mesenchyme cells found in the blastocoel (Burke, 1978; Tamboline and Burke, 1992). Circumesophageal muscle cells originate from SMCs and will eventually form the cicrumesophageal muscles (Burke and Alvarez, 1988). Lastly, coelomic pouch cells develop from SMC and small micromere descendents (Gustafson and Wolpert, 1963). The coelomic pouch cells will form coeloms as outpocketings from the archenteron tip during secondary invagination. The left coelomic pouch will eventually help give rise to the tissue that forms the early adult rudiment.

Isolation of Genes Activated During Pigment Cell Specification

A macroarray library resource for the sea urchin was built by sequencing the ends of 76,020 bacterial artificial chromosome (BAC) recombinants (Cameron et al., 2000). Collectively, several stage and tissue specific libraries were constructed, each consisting of $>10^5$ clones in order to contain representative clones for all genes (Cameron et al., 2000). The purpose of these large cDNA libraries was for gene discovery.

A differential screening of a 20 hour cDNA macroarray library from sea urchin embryos isolated genes activated during pigment cell specification and differentiation (Calestani et al., 2003). A differential screening compares detected signals from two non-specific hybridization probes; one probe is created from a mRNA population including the genes of interest while the other probe is created from a mRNA population missing them. The screening done by Calestani et al. (2003) involved the comparison of LiCl-treated embryos, which create an expansion of the

endomesodermal territory (Hörstadius, 1973, Ransick et al., 1993; Sherwood and McClay, 1997), and dominant negative Notch expressing embryos (lacks Notch intracellular domain), which blocked the signaling between micromeres to macromeres resulting in a loss of SMC specification (Sherwood and McClay, 1999).

A subtractive hybridization was also used in order to identify low-prevalence transcripts (Rast et al., 2000). This was accomplished using a selectate from the LiCl-treated embryos (mRNA population containing the SMC specific transcripts) and a driver from dnN-injected embryos which lack those sequences of interest (SMC specific). This resulted in isolation of genes involved in pigment cell specification because the other three SMC cell types are not completely absent in dnN injected embryos used in the driver. Approximately 400 putative clones were identified and sequenced from the 5' end. A subset of 66 clones underwent gene expression analyses that identified 40 positive clones.

The cDNA clones identified contained sequences highly similar to the following: the transcription factor glial cells missing (*Gcm*), the polyketide synthase gene cluster (*Pks*), three different members of the flavin-containing monooxygenase gene family (*Fmo*), and a sulfotransferase gene (*Sult*) (Calestani et al., 2003). Whole mount in situ hybridization illustrated that the genes are specifically expressed in pigment cells (Calestani et al., 2003) [Figure 5]. At hatched blastula, the genes were expressed in SMC precursors at the vegetal plate. At mid-gastrula, the genes were expressed in cells embedded into the ectoderm. By 72 hours, the genes were still expressed in cells embedded into the ectoderm.
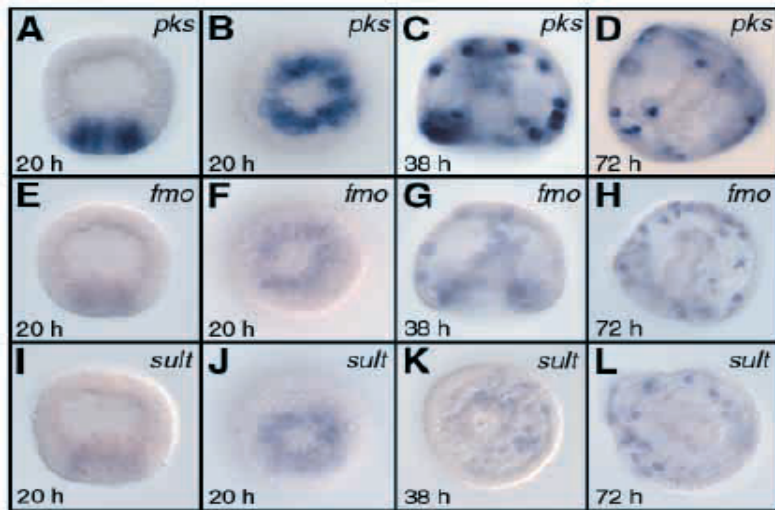
Figure 5: Whole mount *in situ* hybridizaiton of pigment cell genes from hatched blastula through pluteus stage.

Blastula side view (A,E, I), vegetal view (B, F, J), gastrula side view (C and G), apical view (K), and pluteus oral view (D,H,L) [Calestani et al. 2003].

A functional analysis of *SpPks* and *Spfmo1* using antisense technology showed that their expression is essential for the biosynthesis of echinochrome (Calestani et al., 2003) [Figure 6]. A knockdown was done on *Pks* and *Fmo1* resulting in the appearance of morphologically normal embryos, however the embryos did not display any pigment. It was not known whether the results were due to a block of pigment biosynthesis or an absence of pigment cells. Whole mount *in situ* hybridization with an antisense labeled probe for *SpSult* showed that pigment cells were indeed present, and that *Pks* and *Fmo1* are required for echinochrome biosynthesis.
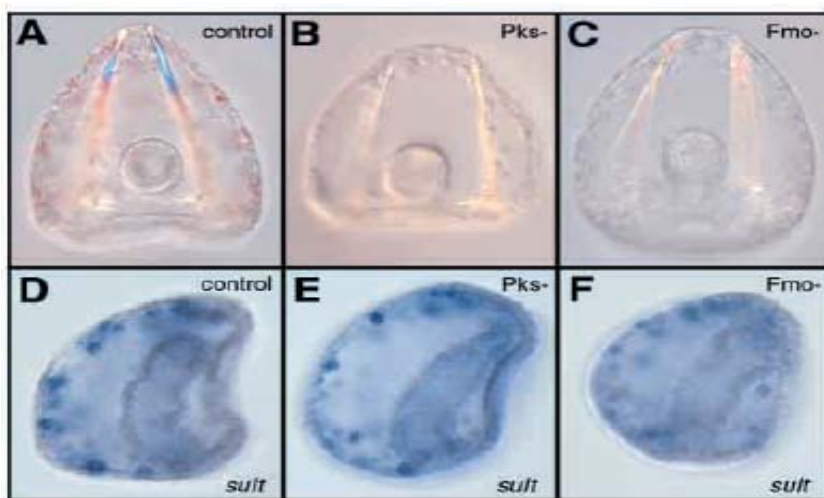
Figure 6: Antisense technology injected with morpholino oligonucleotides (A-C).

Aboral view of prism stage embryos injected with Antisense morpholino oligonucleotides against a random sequence (A), *SpPks* (B), and *SpFmo1* (C). Whole mount in situ hybridizaton confirms that pigment cells are present and that antisense blocks pigment cell biosynthesis (D-F). Aboral view of prism stage embryos injected with Antisense RNA DIG-labeled probe for *SpSult* with a random sequence morpholino oligonucleotide (D), with *SpPks* morpholino olignonucleotide (E), and with *SpFmo1* morpholino oligonucleotide (F) [Calestani et al. 2003].

Polyketide Synthases

Polyketide synthases (PKSs) are a large family of multifunctional proteins mainly found in bacteria, fungi, and plants. They are responsible for the biosynthesis of a wide variety of polyketide compounds (reviewed in Hopwood, 1997; Hopwood and Sherman, 1990; Stuanton and Weissman, 2001). For instance, in bacteria and fungi, PKSs catalyze the biosynthesis of

several antibiotics and mycotoxins. In plants, they synthesize several kinds of compounds ranging from flower pigments to plant-pathogen response compounds (Bohm, 1998).

PKSs are typically divided into three groups, type I, II and III. Type I PKSs are composed of one or more large multifunctional proteins that contain all of the enzymatic activities mandatory for polyketide biosynthesis: acyl transferase (AT), acyl carrier protein (ACP), β-ketoacyl synthase (KS), β-keto reductases (KR), dehydratases (DH), enoyl reductases (ER), and thiolesterase (TE) [Hopwood, 1997; Staunton and Weissman, 2001]. Type II PKSs are composed of numerous monofunctional subunits in which the active sites are distributed among several polypeptides (Hopwood, 1997; Shen, 2000; Staunton and Weissman 2001; Khosla et al. 1999; Keating and Walsh 1999; Hutchinson 1999). Type III PKSs, also known as chalcone synthases, were first isolated in plants and have recently been identified in several bacteria (Moore and Hopke, 2001; Gross et al., 2006). Unlike the first two groups, type III PKSs do not posses a PP domain so polyketide biosynthesis occurs at a single multifunctional active site. Despite functional and structural varieties between the three types, all PKSs biosynthesize the polyketide carbon backbone by successive decarboxylative condensation of simple carbon units (acetyl-CoA and malonyl-CoA) in order to construct a β-keto chain (Kwon et al., 2002).

PKSs are very similar to fatty acid synthases (FASs) in terms of protein domains and biosynthetic pathways. Both PKSs and FASs are responsible for catalyzing successive rounds of decarboxylative condensations of simple carbon units in order to build a β-keto chain. However, during fatty acid biosynthesis, the keto groups are fully reduced as the saturated carbon chains are produced, while in polyketide biosynthesis the keto groups remain partially or unreduced.

In sea urchin embryos, *SpPks* is expressed in pigment cells beginning at the hatched blastula stage and it is maintained throughout the pluteus stage (Calestani et al., 2003). Sequence

analysis identified the following conserved domains: polyketide synthase (PKS) domain, acyltransferase (AT) domain, dehydratase (DH), a zinc-dependent alcohol dehydrogenase (ADH-zinc) domain, and an ACP domain (Calestani et al., 2003; Castoe et al., 2007).

Initial BLAST results of SpPKS against the NCBI database revealed low sequence similarity to FASs in human, *Drosophila*, and *Caenorhabditis* and failed to reveal any other obvious orthologues in other animal genomes. The initial comparisons did show, however, high sequence similarity with fungal and bacterial sequences. Phylogenetic analyses suggested that *SpPks* is most closely related to a *Pks* found in the slime mold *Dictyostelium discoidium* (Castoe et al., 2007).

Sulfotransferases

Sulfotransferases catalyze the sulfate conjugation of several substrates found in bacteria through higher eukaryotes. The catalysis results either in compound detoxification or bioactivation. In mammals, sulfate conjugation is an important pathway for catalyzing hormones, neurotransmitters, drugs and xenobiotics (Falany, 1997; Weinshilboum et al., 1997). In plants, water soluble plant pigments called flavonols, as well as other endogenous compounds, undergo sulfate conjugation (Varin and lbrahim, 1989). Sulfotransferases illustrate different developmental and tissue-specific gene expression (Her at al., 1997; Dunn and Klaassen, 1998). The sea urchin sulfotransferase gene is expressed solely in pigment cells, suggesting that it could play a specific role in pigment cell specific biosynthetic processes (Calestani et al., 2003).

Flavin-containing monooxygenases

Flavin-containing monooxygenases (FMOs) are NADPH-dependent flavoproteins mainly found in bacteria, plants, and higher metazoa. They are responsible for catalyzing the oxidation of several compounds including nucloephilic nitrogen, sulphur, phosphor and selenium atoms. FMOs are primarily involved in the detoxification of various xenobiotics, as well as the molecular activation of numerous metabolites. Three different sea urchin *Fmo* genes have been isolated that may be involved in echinochrome biosynthesis (Calestani et al., 2003). In mammals, five *Fmo* have been isolated thus far (Hines at al., 1994; Gasser, 1996; Cashman, 2000).The three sea urchin *Fmo* genes are expressed in the same cell type (pigment cells) and their expression begins at the same time. Together these results suggest that the three *SpFmo* genes may be co-regulated.

Glial Cells Missing

SpGcm is a sea urchin transcription factor named after one of the two *Drosophila glial cells missing* genes, *gcm1* (Akiyama et al., 1996). The fly *gcm1 and gcm2* encode for transcription factors that are responsible for the development of glial cells in the embryonic nervous system (Hosoya et al., 1995; Jones et al., 1995). Mammals also have two *gcm* genes, however they do not seem to play a role in gliogenesis. Instead, *gcm1* is essential for placental development (Anson-Cartwright et al., 2000) and g*cm2* is necessary for the development of parathyroid glands (Gunther et al., 2000).

Fly *gcm1* and *gcm2*, along with mammalian homolgues, collectively form a small family of transcription factors. These proteins are characterized by a highly conserved domain located in the amino-terminal region, also known as the "gcm box," and comprise the DNA binding site

(Akiyama et al., 1996: Shreiber et al., 1997). In the sea urchin, only one *SpGcm* has been isolated thus far. *SpGcm* is directly activated in SMC precursors at early blastula by Notch signaling (Ransick and Davidson, 2006). It is also known that *SpGcm* positively regulates the expression of several pigment cell specific genes: *SpPks*, *SpSult*, *SpFmo1*, *SpFmo2*, and *SpFmo3* (Davidson et al., 2003). However, it is not known if these are the only genes regulated by *SpGcm* in sea urchin embryos.

Phylogeny of the Sea Urchin

Sea urchins are classified in the phylum Echinodermata which also includes sea cucumbers, starfish, brittle stars, and crinoids. Echinoderms are distinguished from other metaozoans by four synapomorphies (Table 1). Echinoderms and hemichordates are the sister group and closest known relatives to chordates, a phylum of animals classified by the presence of a notochord or spinal column (Figure 7).

Table 1. These synapomorphies distinguish echinoderms from other metazoans (Wray, 1999).

Synapomorphy	Brief Description
Calcitic Skeleton	Endoskeleton composed of calcitic ossicles
Water Vascular System	Functions in movement, respiration, and digestion
"Locked" and "Unlocked" Collagenous Tissue	Ossicles are connected by collagenous ligaments that are mutable
Pentaradial Body Organization in Adults	Adults have a five-fold radial symmetry while larvae are bilaterally symmetrical

Sea urchins are deuterostomes, a group of organisms distinguished by the blastopore forming the larval anus while the mouth is a secondary opening. Previous molecular phylogenetic studies have suggested that deuterostomes only include the vertebrates, the related invertebrate chordates (cephalochordates and urochordates), and three other invertebrate taxa: hemichordates, echinoderms and *Xenoturbella* (Bourlat et al., 2006). Recent phylogenetic analyses have suggested that urochordates, rather than the traditionally accepted cephalochordates, may be the closest invertebrate sister group of vertebrates (Delsuc et al., 2006; Oda et al., 2002; Wada et al., 2006; Jeffery et al., 2004). Even more stunning is the proposition that cephalochordates are more closely related to echinoderms than to vertebrates and urochochordates suggesting that chordates are paraphyletic (Delsuc et al., 2006).

An attractive feature of the sea urchin is its' phylogenetic placement. Along with the hemichordates, sea urchins comprise the sister clade to the chordates within the deuterostomes (Figure 7). In order to fully understand what makes a vertebrate different from other animals, the sea urchin has become a well-studied organism for comparative genomics as a pre-vertebrate starting point.
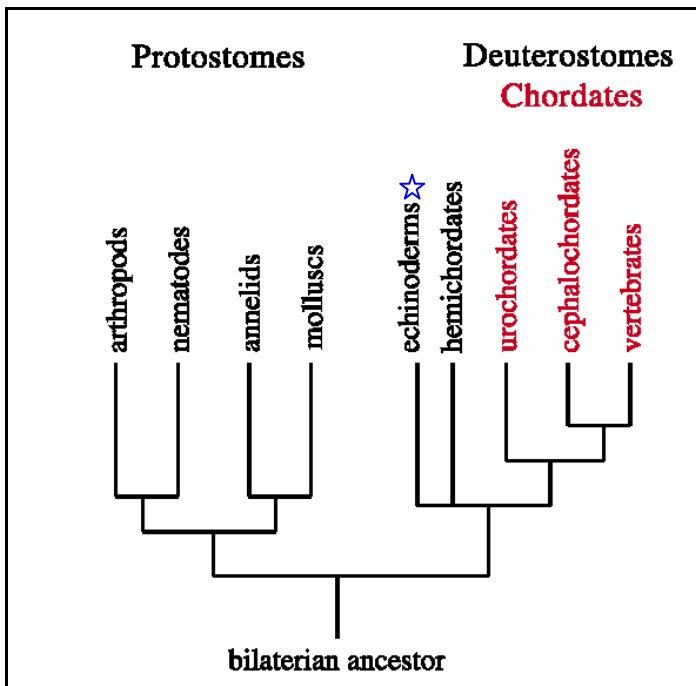
Figure 7: A phylogeny of bilaterians.

The phylogeny depicts the positional status of the sea urchin (indicated by the blue star) in regards to other metazoans (Dehal et al., 2002).

MATERIALS AND METHODS

Pigment Cell Gene Annotation

SpPks

At the time this annotation was performed for *SpPks,* the assembled sea urchin genome sequence was not available. Pigment cell gene annotation of the polyketide synthase gene was carried out using a BAC clone containing the *SpPks* gene. The cDNA contigs were built from 30 cDNA 5' ends sequences for *SpPks* (Calestani et al., 2003). This yielded three sea urchin *Pks* cDNA contigs. The three contigs ranged in size from 938 to 4,232 base pairs in length. Each contig was loaded into the Cartwheel Server (http://cartwheel.caltech.edu/) and a pairwise analysis was conducted between the cDNA contig and BAC sequence Sp_08H21_L. The analysis was then viewed in Family Relations II (http://family.caltech.edu/) which allowed the visualization of where the cDNA contig was located on the BAC sequence.

The sequence annotator program Sea Urchin Gene AnnotatoR (SUGAR) was used to annotate the BAC sequence. Exon predictions were performed using three different software packages (HMMGene, GenScan, GeneID). The coding sequence was translated into protein using the ExPASy translation tool (http://au.expasy.org/tools/dna.html). From there a Conserved Domain Search with Reverse Position Specific BLAST was used to identify conserved domains.

SpFmo1-3 and SpSult

For each cDNA contig, a megaBLAST search was conducted (comparing highly related nucleotide sequences) against the *S. purpuratus* genome to see where the cDNA contigs were

located. The genomic regions and cDNA contigs were loaded into the Cartwheel Server and a single-sequence analysis was conducted based on the genomics regions for each gene. The cDNA contigs and exon predictions were compared to the genomic regions and viewed in Family Relations II.

For each cDNA contig, a blastn (nucleotide vs. nucleotide) search was conducted against the NCBI database in order to recover a full length cDNA sequence. The record for each gene also included the translated protein sequence. From there a Conserved Domain Search with Reverse Position Specific BLAST was used to identify conserved domains for each sequence.

Phylogenetic Analyses

Blast searches were used to compile sequences related to *SpFmo1-3* and *SpSult.* Each pigment cell specific sequence was compared to the genome projects of representative organisms from bacteria to metazoan in order to identify putative orthologues (Tables 2 – 6). The protein sequence was aligned using ClustalX1.81. Phylogenetic analyses were performed with PAUP*4.0. The tree topology was determined through both distance (Neighbor-joining) and character based (Maximum Parsimony) methods.

Neighbor-joining trees were generated using the default settings for the neighbor-joining algorithm and was assessed with 500 bootstrap replications. In order to root the trees, bacteria was defined as the outgroup assuming that bacteria would be the most basal taxon. Maximum Parsimony analysis was obtained by a strict consensus of a full heuristic search with 200 random repetitions of stepwise additions. The confidence at each node was assessed with 500 bootstrap replications.

Table 2: List of SpSULT taxon sequences used in alignment as well as their corresponding GI and Accession Number.

Species	GI Number	Accession Number
Anopheles gambiae	55244326	EAA05552
Bos taurus	76629434	XP_602209
Caenorhabditis elegans	17566452	NP_507880
Canis familiaris	57092833	XP_531772
Crocosphaera watsonii	67852595	EAM48023
Danio rerio	68392177	XP_700097
Drosophila melanogaster	24645281	NP_649870
Gallus gallus	50744552	XP_419772
Gibberella zeae	42549544	EAA72387
Homo sapiens	83754960	2ETG_B
Mus musculus	18490573	AAH22665
Nitrosococcus oceani	77163623	YP_342148
Rattus norvegicus	62654845	XP_346008
Sphingopyxis alaskensis	68524935	EAN48057
Strongylocentrotus purpuratus	76364248	ABA41638
Trichodesmium erythraeum	71669865	EAO26536
Xenopus laevis	27695131	AAH43790
Xenopus tropicalis	58332840	NP_001011496

Table 3: List of SpFMO1 taxon sequences used in alignment as well as their corresponding GI and Accession

Number.

Species	GI Number	Accession Number
Aspergillus fumigatus	66845633	EAL85967
Bos taurus	76637349	XP_580516
Caenorhabditis elegans	25150462	NP_501968
Caenorhabditis briggsae	39581865	CAE60758
Canis familiaris	73961376	XP_547467
Danio rerio	56207486	CAI21026
Drosophila melanogaster	19921694	NP_610217
Gallus gallus	45383027	NP_989910
Gibberella zeae	42552493	EAA75336
Gloeobacter violaceus	37521533	NP_924910
Homo sapiens	6225373	Q99518
Janibacter sp	84384229	EAQ00109
Legionella pneumophila	54293806	YP_126221
Mus musculus	74143660	BAE28877
Rattus norvegicus	62659523	XP_579556
Streptomyces coelicolor	32141333	NP_733734
Strongylocentrotus purpuratus	72124731	XP_791122
Xenopus laevis	50924780	AAH79705
Xenopus tropicalis	49900181	AAH76928

Table 4: List of SpFMO2 taxon sequences used in alignment as well as their corresponding GI and Accession Number.

Species	GI Number	Accession Number
Bos taurus	76637349	XP_580516
Caenorhabditis elegans	17555726	NP_499356
Canis familiaris	57089003	XP_537197
Danio rerio	56207486	CAI21026
Drosophila melanogaster	19921694	NP_610217
Gallus gallus	45383027	NP_989910
Gibberella zeae	42552493	EAA75336.
Gloeobacter violaceus	37521533	NP_924910
Homo sapiens	6225373	Q99518
Legionella pneumophila	54293806	YP_126221
Mus musculus	1899255	AAB50013
Rattus norvegicus	78214354	NP_653338
Rattus rattus	21311522	AAM46763
Saccharomyces cerevisiae	41629686	NP_012046
Streptomyces coelicolor	32141333	NP_733734
Strongylocentrotus purpuratus	72138571	XP_793151
Synechococcus sp	86168955	EAQ70211
Takifugu rubripes	74136131	NP_001027928
Xenopus laevis	50924780	AAH79705
Xenopus laevis	37747791	AAH59977

Table 5: List of SpFMO3 taxon sequences used in alignment as well as their corresponding GI and Accession Number.

Species	GI Number	Accession Number
Bos taurus	76637339	XP_884617
Brevibacterium linens	62426099	ZP_00381228
Caenorhabditis elegans	25150462	NP_501968
Canis familiaris	73961373	CAF74915
Danio rerio	37595430	AAQ94601
Drosophila melanogaster	19922866	NP_611859
Gallus gallus	50744552	XP_419772
Gibberella zeae	42552493	EAA75336
Gloeobacter violaceus	37521533	NP_924910
Homo sapiens	6225373	Q99518|
Janibacter sp	84384229	EAQ00109
Legionella_pneumophila	54293806	YP_126221
Mus musculus	21450117	NP_659127
Neurospora crassa	28924587	EAA33706
Rattus norvegicus	47480111	AAH70904
Rattus rattus	21311522	AAM46763
Streptomyces coelicolor	32141333	NP_733734
Strongylocentrotus purpuratus	72078616	XP_794372
Takifugu rubripes	45502122	CAF74915
Xenopus laevis	50924780	AAH79705
Xenopus tropicalis	62751488	NP_001015707

Table 6: List of SpFMO1-3 taxon sequences used in the alignment as well as their corresponding protein name and accession number.

Species	Protein Name	Accession Number
Caenorhabditis elegans	FMO1	NP_501968
Caenorhabditis elegans	FMO2	NP_501972
Caenorhabditis elegans	FMO3	NP_499356
Caenorhabditis elegans	FMO4	NP_506370
Caenorhabditis elegans	FMO5	NP_503352
Danio rerio	FMO5	NP_944592
Drosophila melanogaster	FMO1	NP_611859
Drosophila melanogaster	FMO2	NP_610217
Gibberella zeae	FMO	EAA75336
Gloeobacter violaceus	FMO	NP_924910.1
Homo sapiens	FMO1	NP_002012
Homo sapiens	FMO2	NP_001451
Homo sapiens	FMO3	NP_001002294
Homo sapiens	FMO4	NP_002013
Homo sapiens	FMO5	NP_001452
Mus musculus	FMO1	NP_034361
Mus musculus	FMO2	NP_061369
Mus musculus	FMO3	NP_032056
Mus musculus	FMO4	NP_659127
Mus musculus	FMO5	NP_034362
Strongylocentrotus purpuratus	FMO1	XP_791122
Strongylocentrotus purpuratus	FMO2	XP_793151
Strongylocentrotus purpuratus	FMO3	XP_794372
Takifugu rubripes	FMO	NP_001027928
Xenopus tropicalis	FMO3	NP_001025595

Procurement of Gametes

Adult sea urchins were rinsed free of any debris with sea water. Spawning was induced (both physically and chemically) and the gametes were viewed in order to distinguish between

males and females (eggs are yellowish-orange in color and sperm are white). During the breeding season (mid November-March) it was possible to induce spawning by vigorous shaking. However, when the physical shaking did not result in gametes, the adult urchins were injected with 2-3mL of 0.5M KCl into different spots around the mouth to chemically induce spawning.

For collection of sperm, males were placed oral side down and the sperm was collected "dry" by pipetting it directly from the gonopores. The sperm was immediately transferred to a microcentrifuge tube and kept on ice. For collection of eggs, females were placed upside down (oral side up) on a small glass beaker of filtered sea water (FSW) placed in an ice bucket. Once settled, eggs were transferred into a microcentrifuge tube and placed in an ice bucket.

Embryo Culture

Twenty-five microliters of eggs were fertilized in petri dishes filled with 25mL of FSW (final concentration of 1000 eggs/mL). Eggs were fertilized in 10mM para-aminobenzoic acid (PABA) sea water. The fertilization envelope that surrounds an embryo is stabilized by the chemical cross-linking of tyrosine residues by ovoperoxidase (Ettensohn, 2004). If ovoperoxidase is blocked, the envelope still forms but it does not stabilize and can be easily removed post-fertilization. PABA sea water inhibits the activity of ovoperoxidase. Approximately one inch of a glass pasteur of sperm was diluted in 10mL of FSW. Dilution activates sperm which have a limited life span and were used within 15 minutes of activation. Approximately 1mL of sperm was added to each petri dish. Once fertilization was complete, the embryos were passed through a 50 micron Nitex nylon mesh in order to remove the fertilization

envelope. Penicillin (20 units/ml) and streptomycin (50 ug/ml) was added to each petri dish to repress bacterial growth.

Within the first few hours of development, embryos were transferred to fresh FSW in order to promote a healthy culture. The development of the culture was monitored to ensure that the embryos were developing normally. Embryo cultures were kept in an incubator set at 16° C throughout development.

RNA Extraction

Embryos for each time point were homogenized in 800μL RNA-Bee, a solution containing phenol and guanidine thiocyanate. One hundred sixty microliters of choloroform was added to each homogenate, vortexed for 15 seconds and stored on ice for 5 minutes. The homogenate was then centrifuged at 13,200 rpm for 15 minutes at 4° C. After centrifugation, the aqueous phase was transferred into a new microcentrifuge tube (RNA resides in the aqueous phase while DNA and proteins remain in the organic phase and interphase). Five micrograms of glycogen and 0.5mL of isopropanol was added to each sample. It was stored at room temperature for 10 minutes and then centrifuged at 13,200 rpm for 5 minutes at 4° C. The supernatant was removed from the sample and the RNA pellet was washed with 1mL of 75% ethanol. The sample was then centrifuged for 5 minutes at 9,600 rpm at 4° C. Following centrifugation, the supernatant was removed and the RNA pellet was air-dried. The RNA was dissolved in 25μL of nuclease-free water. Dissolved RNA was stored at -80° C.

Gene Expression Profile

Quantitative real-time PCR (QPCR) was used to measure the relative amount of *SpPks*, *SpFmo1-3, SpSult* and *SpGcm* transcripts during the course of development (egg, 8h, 12h, 15h, 18h, 21h, 24h, 27h, 30h, 40h, 50, 62h, and 72h). The total RNA isolated from embryos at each stage was DNase treated (DNA-free, Ambion). For each sample, 0.1 volume of 10X DNase I Buffer (final concentration) and 2 units of rDNase I were added to the RNA. The samples were incubated for 20-30 minutes at 37° C. First strand cDNA was synthesized by reverse transcription using 1μg of RNA (High Capacity cDNA Archive Kit, Applied Biosystems). Reverse transcription reactions were prepared according to the manufacturer (total reaction volume was 50 ul; reaction composition was: 10μL 10XRT buffer, 4μL 25X dNTP mixture, 10μL 10X Random Primers, 50U of MultiScribe RT).

First strand cDNA was used as the template for the QPCR reactions prepared as followed: 1μL of cDNA, 15pmol of Forward and Reverse primers, 15μL Sybr® Green PCR Master Mix, and 12μL nuclease free water. The reactions were run in triplicate for each developmental stage/primer set combination on an ABI PRISM 7000 Sequence Detection System. A non-template control for each primer set and a non-RT control (DNase treated RNA as a template) for each developmental stage were included.

QPCR is a method that allows quantification of starting amounts of cDNA templates. Overall, the methodology is based on a fluorescent reporter molecule, such as Sybr Green that binds to double-stranded DNA. With each cycle, more PCR product is amplified and the detection of the fluorescence increases. During the first few cycles of QPCR there is very little change in fluorescence detection and these stages identify the baseline. An increased detection of fluorescence above the baseline signifies an increase in PCR product. Above the baseline is a

threshold cycle (Ct), a fixed fractional cycle number at which the fluorescence is to be at the most exponential phase possible. Samples were run in triplicates and Ct values were averaged. The Ct average for each primer set was normalized to the ubiquitin Ct for each reaction. Ubiquitin expression is known to be approximately constant throughout development (Nemer et al., 1991 ; Ransick et al., 2002).

<u>Northern Blotting</u>

RNA isolation

Embryos were cultured as previously described and were collected for 25 and 40 hour stages. RNA was isolated using the RNA-Bee protocol as previously described.

Denaturing Gel Agarose Electrophoresis

A 1% denaturing gel was prepared (75mL DEPC water, 1g agarose, 10mL of MOPS 10X buffer (0.2M MOPS pH7, 20mM sodium acetate, 10mM EDTA pH 8), and 15mL 37% Formaldehyde) in 1X MOPS buffer. Twenty-five micrograms for each stage of RNA was combined with loading buffer (1X final concentration) and heated at 65° C for five minutes. The samples were put immediately on ice and 1μL of diluted Ethidium Bromide (EtBr) was added to each tube. A RNA high molecular weight ladder (0.5, 1, 1.5, 2, 3, 4, 6, 8, and 10kb) was prepared using 10μL ladder and 10μL 2X loading buffer. It was heated at 65° C for five minutes, put immediately on ice, and 1μL of diluted EtBr was added. The samples were run at 4volts/cm.

Probe Generation

DNA probes about 500 bp in length were synthesized using standard PCR (10ng of plasmid DNA for each gene, 10pmol of both Forward and Reverse primers, 5μL 10X Buffer plus MgCl$_2$, 1.25U of Taq polymerase, in a total reaction volume of 50μL). The PCR was purified using Promega PCR Clean-Up System. Five microliters of each sample (5μL sample and 1μL 6X loading dye) was run on a 1.2% agarose gel (0.6g agarose, 49.5mL Tris Boric Acid EDTA (TBE) (89mM Tris, 89mM boric adic, 2mM EDTA pH 8.4), and 2μL EtBr) in TBE buffer for verification of probe size.

The probe was radioactively labeled using the Promega Prime-a-Gene® Labeling System. The DNA probes were denatured at 95° C for two minutes and put immediately on ice. The reaction consisted of the following: 10μL Labeling 5X Buffer, 2μL of unlabeled dNTPs (20μM each of dGTP, dCTP, and dTTP), 25ng denatured template, 2μL of Nuclease-Free BSA, 5 units of DNA Polymerase 1 Large (Klenow) Fragment, 5μL of [α-^{32} P] dATP (50μCi), and nuclease free water to bring the final volume to 50μL. The samples sat at room temperature for one hour.

Transfer to Solid Support and Immobilization

The gel was rinsed several times in DEPC water to remove excess formaldehyde. A corner of the gel was cut off for orientation (the corner of the first loaded well). A piece of Millipore Immobilon-Ny+ Transfer Membrane was cut to the size of the agarose gel. The membrane was hydrated with DEPC water and transferred into transfer buffer (20X SSC). A glass tray was filled with 500mL of transfer solution (20X SSC) and a glass plate was laid on top as a support. Two sheets of Whatman® 3MM filter paper was soaked in 20X SSC and the sheets

were laid across the glass support with the ends placed in the 20X SCC in the glass tray. The gel was flipped over and placed directly on the filter paper. The membrane was then placed on top of the gel. Parafilm was placed along the edge of the gel to avoid the flow of buffer around the borders of the gel. Three sheets of Whatman® 3MM filter paper were cut to the size of the gel and soaked in 20X SSC. They were placed on top of the membrane and air bubbles were removed. A stack of absorbent napkins (approximately 20cm) and a weight of approximately 300g were placed on top. The RNA was left to transfer from the gel to the membrane overnight (approximately 12 to 18 hours).

Once transferred the napkins and filter paper were discarded. The blot was lifted from the gel using forceps and rinsed in 6X SSC to eliminate any loose particles or agarose. The blot was placed on a sheet of filter paper and left overnight to dry. Once transferred the blot was rinsed in 6X SSC and left overnight to air dry.

Prehybridization and Hybridization

The blot was hydrated in DEPC water and then put into a hybridization bottle with the RNA facing the inside of the tube. Next, 100mL of Hybridization Buffer (0.5M Sodium Phosphate pH 7.2, 2mM EDTA, 7% (w/v) SDS) was added to pre-hybridize the blot and was placed into a rotating hybridization oven set at 68° C for 1-2 hours. The probe was denatured with 5μL of 3M NaOH prior to addition to the filters. The hybridization was performed in 100mL of fresh Hybridization Buffer containing the radioactively labeled probe. The hybridization bottles were placed into the rotating hybridization oven set at 68° C overnight (approximately 12-18 hours).

Washing

After hybridization the Hybridization Buffer with the probe was discarded and the

hybridization bottle was rinsed three times with 60mL Wash Solution I (1X SSC, 0.1% SDS) at

room temperature. The blot was washed two times with 100mL Wash Solution I heated to 68° C

for 15 minutes and two times with 100mL Wash Solution II (0.5X SSC, 0.1% SDS) at 68° C for

15 minutes.

Detection

The blot was removed from the hybridization bottle, wrapped in saran wrap, labeled,

placed into the storage phosphor screen cassette (Molecular Dynamics), and left to expose for up

to one week. The phosphor screens were scanned at 200 μm resolution by a Phosphorimager

Storm 820 (Molecular Dynamics). In order to determine transcript length, the EtBr stained ladder

was aligned with the autoradiography image of the Northern blot.

Whole Mount *In Situ* Hybridization (WMISH)

The cDNA for Sp*Pks*, *SpSult*, and *SpFmo1-3* was previously isolated from a random

primed cDNA library constructed into pSPORT1 vector (Invitrogen Life Technologies)

(Cameron et al., 2000). The cDNA clones were amplified in bacteria. Five ng of plasmid were

added to Fusion-Blue Competent Cells (Clontech, laboratories), mixed gently, and left in ice for

30 minutes. The cells were heat shocked for 45 seconds in a water bath at 42º C and then put

immediately on ice. After heat shocking, 450μL of SOC medium was added to each sample and

then incubated in a shaker (250 rpm) at 37º C for one hour. Next, 50μL of the cells for each

transformation was spread on plates containing ampicillin (100ug/ml) and incubated at 37° C over night. The next day, a single colony was picked from each plate, placed in 3mL of LB with ampicillin (100ug/ml), and incubated overnight in a shaker at 37° C.

Plasmid DNA was purified using QIAprep Spin Miniprep Kit (QIAGEN). The bacterial cells were resuspended and 250µL of Buffer P1 was added to each tube. Then, 250µL of Buffer P2 was added to each tube and inverted to mix. Next, 350µL of Buffer N3 was added to each tube and inverted to mix. The samples were centrifuged for 10 minutes at 13,000 rpm. The resulting supernatant was transferred into a QIAprep spin column and centrifuged for one minute. The flow through was discarded and 0.5mL of buffer PB was added to each column and the samples were centrifuged for one minute. Once again the flow through was discarded, 0.75mL Buffer PE was added to each column, and the samples were centrifuged for one minute. The flow through was discarded and the columns were centrifuged for one minute to eliminate excess wash buffer. Lastly, 50µL of Buffer EB (10mM Tris ph 8.5) was added to each column and was centrifuged for one minute to elute DNA.

Prehybridization

Embryos were cultured as previously described and were collected at the following stages: 18, 21, 24, 27, 30, 40, 50, and 63 hours. The fertilization membrane was removed for cleavage stage embryos (pre-hatching) as previously described in order to obtain satisfactory results. Embryos were stored in freshly prepared fixative according to the Harkey protocol (2.5% glutaraldehyde, 0.14M NaCl, 0.2M Sodium-Phosphate buffer pH 7.4) (Harkey, 1992) (Ransick et al., 2002). After each wash or rinse, embryos were spun down for 15 seconds at 1,000 rpm and the supernatant was removed (unless otherwise noted). The fixative was washed out with two

exchanges of ice cold wash buffer (0.3M NaCl, 0.1M Sodium-Phosphate Buffer pH 7.4, 0.1%

Tween 20, DEPC water). Then the embryos were washed three times with 500μL of PBST

(150mM NaCl, 0.1M Sodium-Phosphate Buffer 0.1% tween, DEPC water) and left at room

temperature for 15-20 minutes. Next, 200μL of Proteinase-K in PBST (5μg/mL) was added and

left to incubate at room temperature for five minutes. The embryos were washed two times with

500μL glycine in PBST (2mg/mL) and washed one time with 1mL of PBST.

The embryos were post-fixed with 500μL 4% paraformaldehyde in PBS (150mM NaCl,

0.1M Sodium-Phophate Buffer pH 7.4, DEPC water) at room temperature for 30 minutes. The

embryos were then washed three times with 500μL PBST leaving the last amount inside the

microcentrifuge tube. Hybridization buffer (50% Formamide, 10% PEG, 0.6M NaCl, 5mM

EDTA, 20mM Tris pH 7.5, 2X Denhardts, 0.1% Tween 20, DEPC water) was added in a 1:1

ratio of PBST. The supernatant was removed and 500μL of hybridization was added, mixed

thoroughly, and incubated for one hour at 55° C.

Probe Preparation

Single-stranded RNA probes were synthesized using RNA polymerase (Sp6 and T7) with

the linearized cDNA clone in pSPORT1 as template. The samples were linearized by digesting

separately for each gene with NotI and SalI (10μg plasmid, 4μL NEB3 buffer 10X, 4μL BSA

10X, 80U enzyme, in a total volume reaction of 40μl) at 37° C for two hours. After digestion, a

final concentration of 0.5% SDS and 4 μL of Proteinase K (1mg/mL) was added to each sample

and was incubated at 50° C for at least 30 minutes. An equal volume of phenol/chloroform was

added to each sample, vortexed thoroughly, and centrifuged at full speed for 1 minute. The

aqueous phase was transferred into a new centrifuge tube and 10μg of glyocogen was added to

each sample. Also, 1/10th of the volume of 3M Na acetate pH 5.2 and 2 volumes of EtOH were added to each sample and left to incubate at -20° for at least 15 minutes. The samples were centrifuged at full speed for 15 minutes and the supernatant was removed. The pellet was washed with 1mL of 70% EtOH and air dried. The pellet was then re-suspended in 20µL of DEPC water.

The probes were labeled with digoxygenin-11-UTP (DIG) (14µL of 500ng and DEPC water, 2µL of 10X buffer, 2µL DIG, 1µL RNase inhibitor, 1µL of T7 or Sp6 enzyme) for two hours at 37° C. The RNA product was precipitated by adding LiCl and was incubated at 65° C for 10 minutes. The samples were then DNase treated. In each tube, 0.1 volume 10X DNase I Buffer and 1µL rDNase I was added and mixed gently. The samples were incubated at 37°C for 20-30 minutes.

The DIG labeled RNA probes were spun down at full speed for 20 minutes at 4° C. The supernatant was removed, washed with 75% EtOH and spun down for 5 minutes at 4° C. The samples were air dried and resuspended in 20µL DEPC water. Probes were stored at -80° C.

Hybridization

The probes were diluted to 1ng/uL with DEPC water. The probe was added to hybridization buffer (0.02ng/µL), heated at 95° C for 10 minutes, and then immediately transferred to ice. The supernatant was removed from the embryos and 200µL of the probe in hybridization buffer was added to the samples. The samples were mixed thoroughly and incubated overnight at 65° C.

Post-hybridization Washes

Embryos were rinsed with 200μL of a 1:1 hybridization buffer:PBST at room

temperature and mixed thoroughly. Then the embryos were washed two times with 500μL PBST

at 65° C and washed three times with 500μL 1XSSC (with 0.1% Tween) at 65° C for 20 minutes.

Staining

The embryos were rinsed two times with 500μL PBST at room temperature and then

500μL of PBST containing 5% sheep serum was added to each sample and incubated for 30

minutes at room temperature. Next, 200μL of PBST containing 5% sheep serum with sheep anti-

DIG antibodies conjugated to bacterial alkaline phosphatase (1:500 dilution) was added and the

samples were incubated at room temperature for 1 hour.

The samples were rinsed three times with 500μL of PBST and then rinsed two times with

500μL TBST pH 8 (0.1M Tris pH 8, 0.15M NaCl, 0.1% Tween 20, DEPC water) in order to

remove excess and unbound antibody. The samples were then washed once in alkaline phosphate

(AP) staining buffer (100mL Tris pH 9.5, 100mM NaCl, 50mM $MgCl_2$, 1mM levamisole, 0.1%

Tween 20). The embryos were transferred into a 96-well Falcon plate.

Staining was initiated by adding 100μL of AP buffer containing a staining solution of

histochemical substrates BCIP (5-bromo-4-chloro-3-indolyl phosphate) and NBT (nitro blue

tetra-zolium) into each well. The final staining solution consisted of 1mL AP buffer including

3.5μL BCIP and 4.5μL NBT. The plate was incubated in the dark for up to three hours. The plate

was checked every half hour to halt the staining of embryos as needed. Staining was halted by

removing the staining mix and washing the embryos two times with 100μL PBST containing

1mM EDTA. The stained embryos were transitioned to 50% glycerol (15%, 30%, and 50%) and observed at the optical microscope.

Microinjection of Morpholino Antisense Oligonucleotides

Embryos were injected using PicoSpritzer III (Parker Instrumentation), an air pressure system for delivery of picoliters volumes of solutions. Eggs were first de-jellied by placing them in a petri dish filled with acidic sea water with a pH of 4.75 (approximately 135µL 0.3M of citric acid in 50mL of FSW) for one minute. Eggs were washed in FSW three times. This procedure removes the outer polysaccharide layer from the egg and allows the egg to stick to a protamine sulfate coated petri dish. Prior to egg rowing, dishes were filled with 1% protamine sulfate solution for two minutes, rinsed with deionized water, and air dried on the lab bench. The treated dishes allow for the negatively charged surfaces of the eggs to attract and adhere to the positively charged dishes. Thus causing the eggs to be immobilized which allows for rapid injection without affecting development.

De-jellied eggs (approximately 150) were rowed in straight lines on 1% protamine-sulfate treated 60mm cell culture dish lids. A 250µM microinjection solution was made for each gene of interest (600µM morpholino antisense oligonucleotide or a control morpholino oligonucleotide, corresponding to 2.5µL morpholino, 3µL KCl 0.4M, and 0.5µL water). Injection pipettes were created using a P-97 flaming micropipette puller (Sutter Instrument Company). Approximately 1.5µL of the microinjection solution was put into an injection pipette and attached to the pipette holder connected to an inverted microscope and micromanipulators. In order to break the tip of the micropipette, the tip of the micropipette was carefully pressed against the bottom of the plate. The eggs were fertilized with diluted sperm in PABA SW to

prevent hardening of the fertilization envelope. The embryos were stored in an incubator set at

16° C and were observed at the microscope throughout development.

RESULTS

Gene Structure and Annotation of *SpPks*, *SpSult* and *SpFmo1-3*

One of the main goals of this research is to provide ground work for future studies of gene expression regulation. Cis-regulatory studies can be localized in the 5', 3', and intronic regions of a gene. In order for future studies to advance, all of the genomic regions listed above must be identified for the pigment cell specific genes. At the time this annotation was performed for *SpPks*, the assembled genome sequence was not available. Instead, the gene annotation was carried out using a BAC clone containing the *SpPks* gene. The cDNA contigs were compared to the BAC sequence in order to visualize where these contigs were located on the BAC sequence. By comparison, it was possible to identify missing cDNA sequence using the exon prediction packages of the BAC sequence in SUGAR (Figure 8). The predicted exons and protein sequence of *SpPks* was verified by tiling expressed sequence tag (EST) data from partial cDNA clones covering the corresponding genomic region cloned into a BAC.

The missing cDNA sequence was recovered from the BAC sequence and final length of the transcript was 8,447 base pairs in length with seven exons. The SpPKS included three domains highly similar to: the polyketide synthase domain (KS), an acyltransferase domain (AT), and a zinc-binding dehydrogenase domain (ADH-zinc) [Figure 9].

For *SpSult* and *SpFmo1-3*, each cDNA contig was BLASTed against the *S. purpuratus* genome and full length predicted sequences were recovered. The genome was annotated using GLEAN, a program which combines four different exon prediction packages (Elsik et al.,2006). They are not the same packages used in SUGAR. The cDNA contig for *SpSult* was 1,123 base

pairs in length. The partial sequence was compared to the genome sequence and exon prediction packages in SUGAR. The packages predicted 4-7 exons which corresponded to only some of the cDNA sequence (Figure 10). The *SpSult* transcript length recovered from the genome was 858 base pairs in length containing seven predicted exons. This discrepancy in length is due to the error associated with computational intron-exon predictions. Intron-exon computational predictions are often unable to determine the correct length of the 5' and 3' UTR. The domain analysis showed that SpSULT has high similarity to a sulfotransferase domain (Figure 11).

The partial cDNA contig for *SpFmo1* was 919 base pairs in length. The partial sequence was compared to the genome sequence and exon prediction packages in SUGAR. The packages predicted seven to ten exons (Figure 12). The *SpFmo1* transcript recovered from the genome was 1,949 base pairs in length with seven predicted exons. The domain analysis showed that SpFMO1 has high sequence similarity to three domains: a lysine/ornithine N-monooxygenase domain, a flavin-binding monooxygenase-like domain, and a partial uncharacterized conserved domain (Figure 13).

The partial cDNA contig for *SpFmo2* was 943 base pairs in length. The partial sequence was compared to the genome sequence and exon prediction packages in SUGAR. The packages predicted eight to 14 exons (Figure 14). The *SpFmo2* transcript recovered from the genome was 1,192 base pairs in length with eight predicted exons. The domain analysis showed that SpFMO2 has high sequence similarity to two domains: a flavin-binding monoxygenase-like domain and a partial uncharacterized conserved domain (Figure 15).

The partial cDNA contig sequence for *SpFmo3* was 1,033 base pairs in length. The partial sequence was compared to the genome sequence and exon prediction packages in SUGAR. The packages predicted seven to nine exons (Figure 16). The *SpFmo1* transcript

recovered from the genome was 2,200 base pairs in length containing eight predicted exons. The

domain analysis showed that SpFMO3 has high sequence similarity to a lysine/ornithine N-

monooxygenase domain and flavin-binding monooxygenase-like domain (Figure 17).
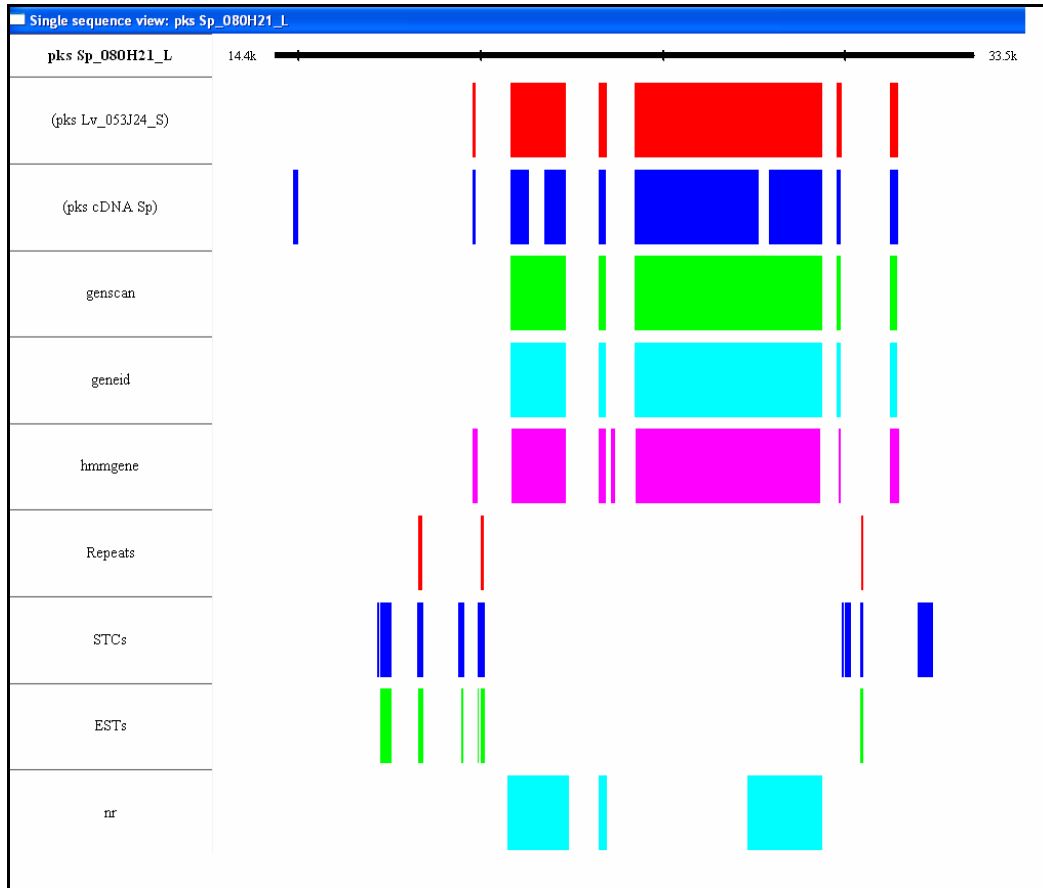
Figure 8: Main screen of the Sea Urchin Gene AnnotatoR for the Sp_08H21_L BAC sequence.

From the top: pks BAC sequence of *S. purpuratus*, pks BAC sequence of *Lytechinus variegatus*, *SpPks*

sequence, three intron/exon identification programs (HMMGene, GenScan, and GeneID), *S. purpuratus*

repeat sequences (Repeats), BAC ends (STCs), sea urchin expressed sequence tags (ESTs), and BLAST

search results against the NCBI non redundant database (nr).
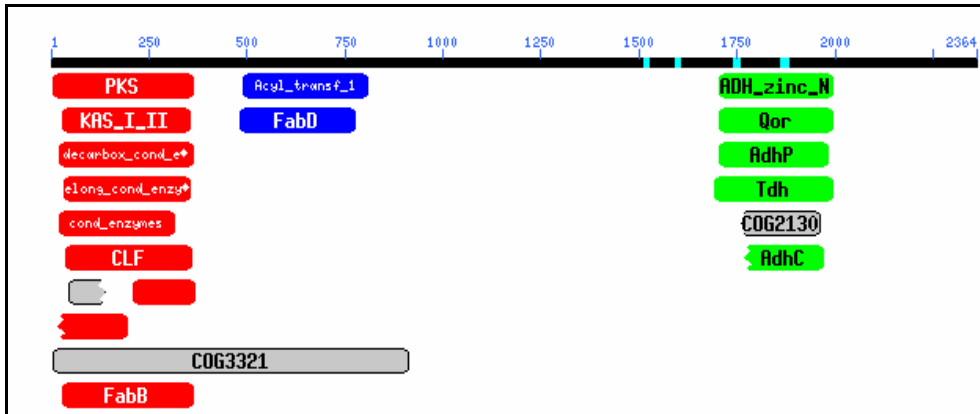
Figure 9: SpPks sequence analysis identified three domains and their associated domains.

The first was a polyketide synthase (PKS) domain, the second was an acyltransferase (AT) domain, and the third was a zinc-dependent alcohol dehydrogenase (ADH-zinc) domain.

Figure 10: *SpSult* cDNA contig and exon predictions were compared to the sea urchin genome sequence.

From the top: Blastn search results against the cDNA contig followed by three intron/exon identification programs (HMMGene, GenScan, and GeneID).

Figure 11: *SpSult* sequence analysis identified high similarity to a sulfotransferase domain.

Figure 12: *SpFmo1* cDNA contig and exon predictions were compared to the sea urchin genome sequence.

From the top: Blastn search results against the cDNA contig followed by three intron/exon identification programs (HMMGene, GenScan, and GeneID).

Figure 13: *SpFmo1* sequence analysis identified three domains.

The first was a lysine/ornithine N-monooxygenase domain (IucD), the second was a flavin-binding monooxygenase-like domain, and the third was a partial uncharacterized conserved domain (COG3349).

50

Figure 14: *SpFmo2* cDNA contig and exon predictions were compared to the sea urchin genome sequence.

From the top: Blastn search results against the cDNA contig followed by three intron/exon identification programs (HMMGene, GenScan, and GeneID).

Figure 15: *SpFmo2* sequence analysis identified two domains.

The first was a flavin-binding monooxygenase-like domain and the second was a partial uncharacterized conserved domain (COG3349).
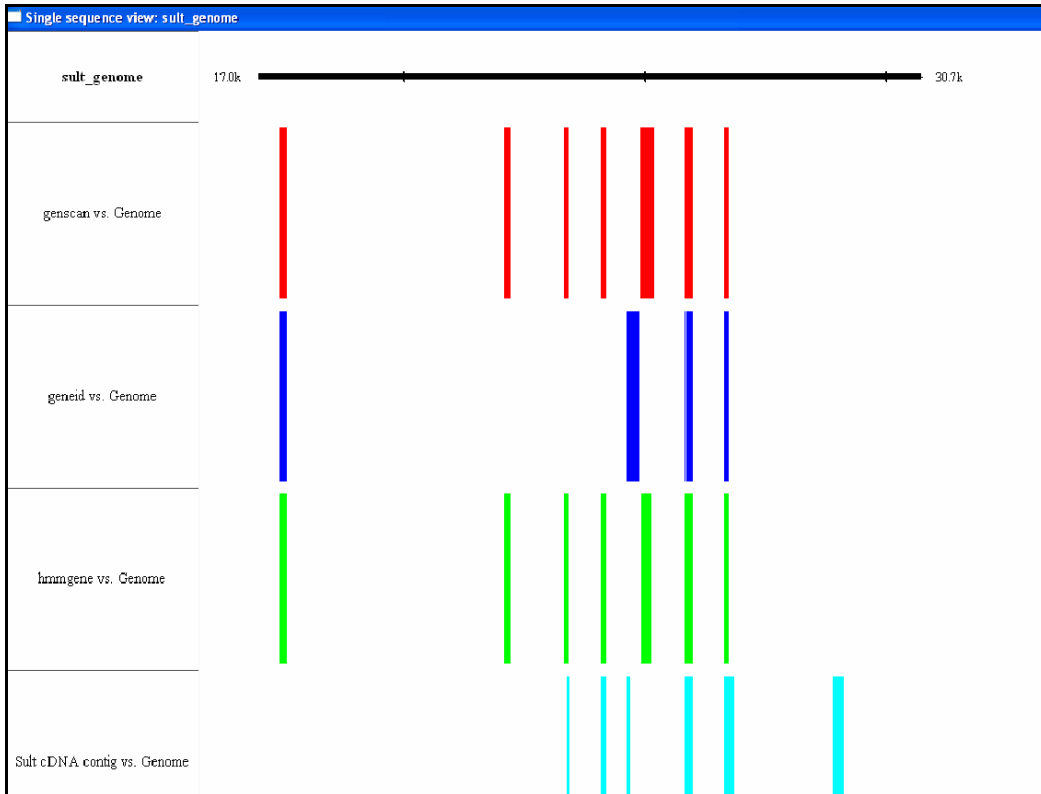
51

Figure 16: *SpFmo3* cDNA contig and exon predictions were compared to the sea urchin genome sequence.

From the top: Blastn search results against the cDNA contig followed by three intron/exon identification programs (HMMGene, GenScan, and GeneID).
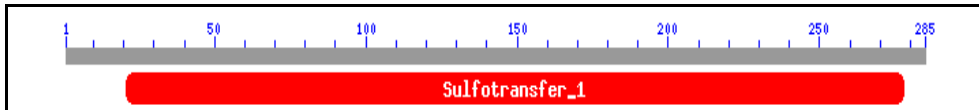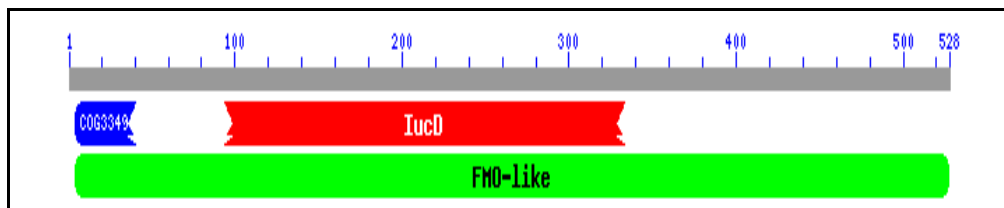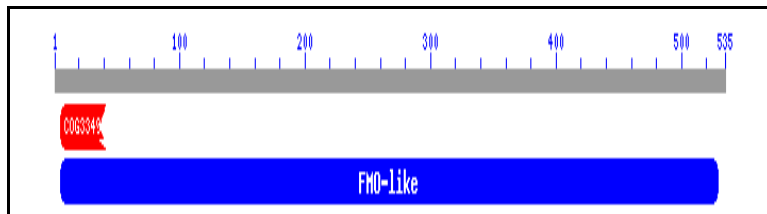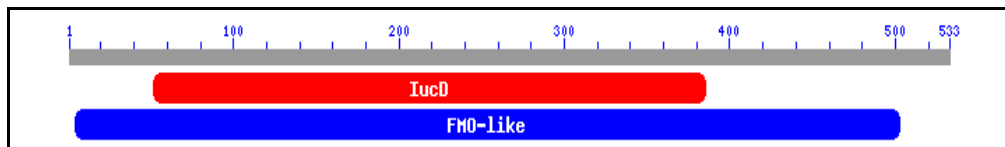

Figure 17: *SpFmo3* sequence analysis identified two domains.

The first was a flavin-binding monooxygenase-like domain and the second was a lysine/ornithine N-monooxygenase domain (IucD).

Phylogenetic Analyses

An attractive feature of the sea urchin is its' phylogenetic placement. In conjunction with the hemichordates, sea urchins comprise the sister clade to the chordates within the deuterostome superphylum (See Figure 7). In order to fully understand what makes a vertebrate different from other animals, the sea urchin has become a well-studied organism for comparative genomics as a pre-vertebrate starting point.

Prior to this research, phylogenetic analyses conducted on *SpPks* suggested a complex evolutionary history of the gene in which it was most closely related to the slime mold *Dictyostelium discoidium pks* (Castoe et al., 2007). Due to the complex evolutionary origins of *SpPks* and lack of animal orthologues, phylogenetic analyses were conducted on *SpSult* and *SpFmo1-3* in order to evaluate the relationships of these other sea urchin pigment cell genes to putative orthologues from other taxa.

Phylogenetic analyses were conducted on *SpSult* and *SpFmo1-3* using both character and distance based approaches. By making use of two different approaches, it is possible to cross-validate phylogenetic estimations since these two approaches have different advancements and assumptions about evolutionary processes. Distance based methods focus on the measurement of distance, for example, the number of amino acid replacements. Neighbor-joining was the method of choice. This particular method utilizes an algorithm that generates a topology with the least total branch length at each step. Although speedy and ideal for large data sets, it does not guarantee true tree topology with the least total branch length due to the fact that it is a "greedy" algorithm that builds the tree in a step-wise manner. Character based methods, on the other hand, focus on a particular state, for example, the presence/absence of an insertion at a particular site. Maximum Parsimony was used as the method of choice. Based on the sequence information, the

method utilizes a matrix of distinct phylogenetic characters in order to generate one or more optimal trees. Unlike Neighbor-joining, the method is much slower and is based solely on informative sites.

Phylogenetic analyses were conducted on *SpSult* and *SpFmo1-3* in order to evaluate the relationships of the sea urchin genes to putative orthologs from other taxa. Bacteria were used as an outgroup for each tree. The tree topology is depicted using both Neighbor-joining and Maximum Parsimony. Neighbor-joining trees were generated by default settings for the neighbor-joining algorithm. Maximum Parsimony trees were generated by a strict consensus of a full heuristic search with 200 random repetitions of stepwise additions. The confidence at each node was assessed with 500 boostrap replications.

For *SpSult*, both the Maximum Parsimony (bootstrap value of 64%) and Neighbor-joining (bootstrap value of 52%) trees suggest that the sea urchin gene is most closely related to vertebrate genes (Figures 18 and 19). For *SpFmo1*, both trees suggest that the sea urchin gene is most closely related to the nematode and vertebrates (bootstrap value of 82% for Maximum Parsimony tree and 100% for Neighbor-joining) [Figures 20 and 21]. The Maximum Parsimony tree for *SpFmo2* suggests that the sea urchin gene is most closely related to fungi, however this placement is weakly supported by bootstrap analyses (bootstrap value < 50%) [Figure 22]. Alternately, the Neighbor-joining tree suggests that *SpFmo2* is most closely related to the nematode and vertebrates (bootstrap value <50%) [Figure 23]. The Maximum Parsimony tree for *SpFmo3* suggests that the sea urchin gene is most closely related to the nematode and vertebrates (bootstrap value <50%) [Figure 24] while the Neighbor-joining places it in a clade with a bacterium, although this placement is weakly supported (bootstrap value of < 50%) [Figure 25]. A gene family tree was put together including the three sea urchin *Fmo* genes, five from human,

five from mouse, one from frog, one from zebrafish, one from pufferfish, five from the

nematode, and two from the fruitfly. Both trees (bootstrap values of 99%) suggest that the three

sea urchin *Fmo* genes are most closely related to each other than to a gene from another

organism (Figures 26 and 27).

Figure 18: The tree topology of *SpSult* was determined through Maximum Parsimony analysis.

The confidence at each node was assessed with 500 bootstrap replications. Vertebrates are highlighted by the red line and the sea urchin gene is indicated by a blue star.

Figure 19: The tree topology of *SpSult* was determined through a Neighbor-joining analysis.

The confidence at each node was assessed with 500 bootstrap replications. Vertebrates are highlighted by the red line and the sea urchin gene is indicated by a blue star.

Figure 20: The tree topology of *SpFmo1* was determined through Maximum Parsimony analysis.

The confidence at each node was assessed with 500 bootstrap replications. Vertebrates are highlighted by

the red line and the sea urchin gene is indicated by a blue star.

Figure 21: The tree topology of *SpFmo1* was determined through a Neighbor-joining analysis.

The confidence at each node was assessed with 500 bootstrap replications. Vertebrates are highlighted by the red line and the sea urchin gene is indicated by a blue star.
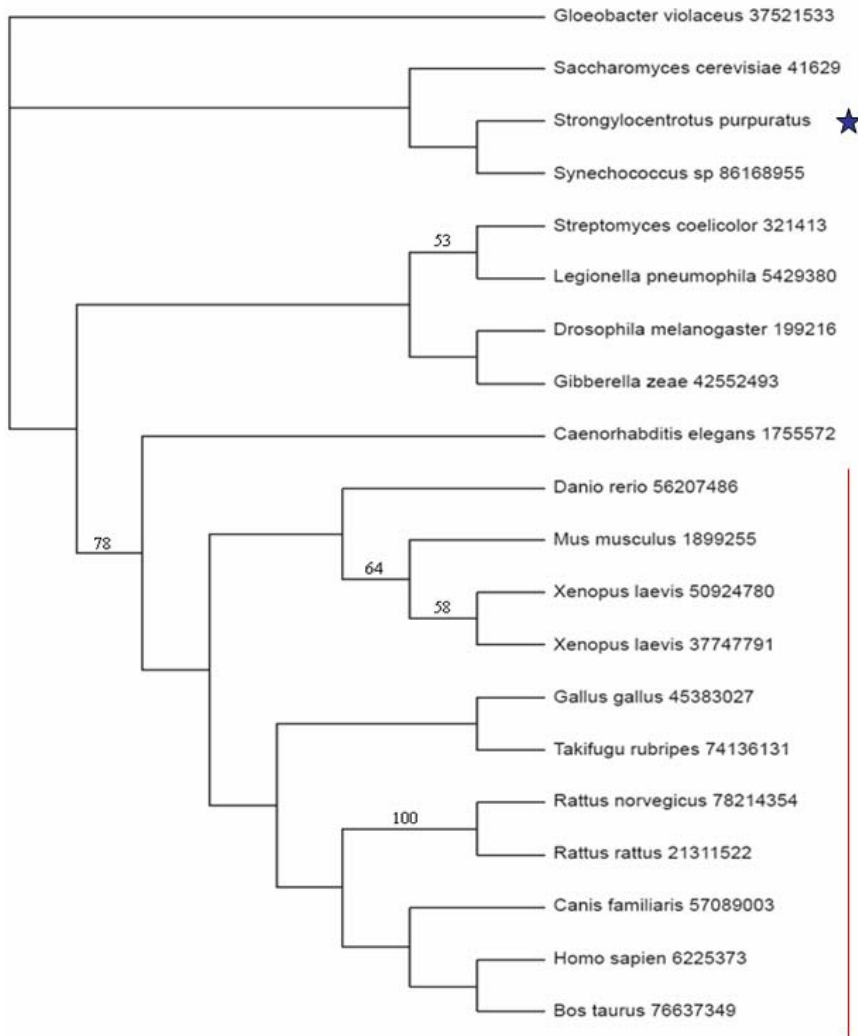
Figure 22: The tree topology of SpFmo2 was determined through Maximum Parsimony analysis.

The confidence at each node was assessed with 500 bootstrap replications. Vertebrates are highlighted by the red line and the sea urchin gene is indicated by a blue star.
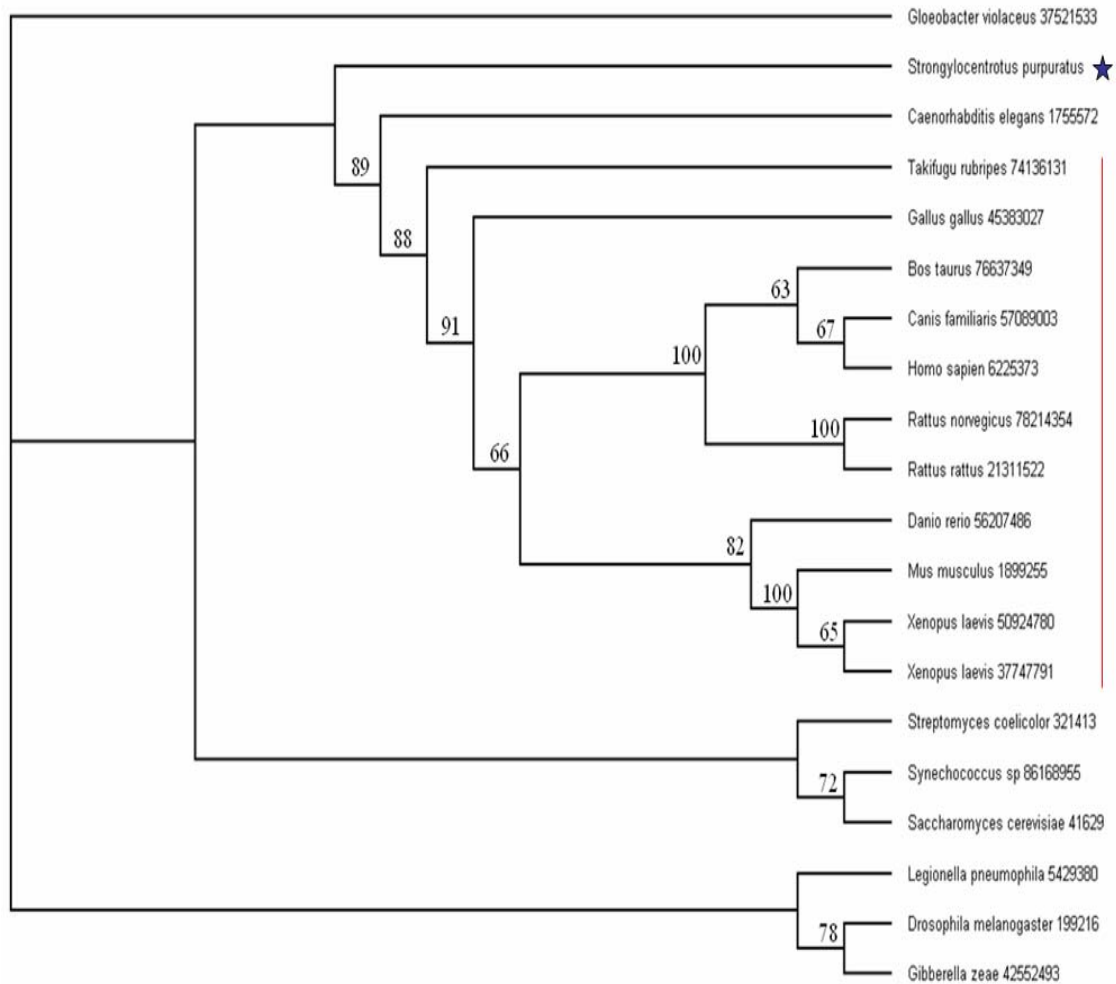
Figure 23: The tree topology of *SpFmo2* was determined through a Neighbor-joining analysis.

The confidence at each node was assessed with 500 bootstrap replications. Vertebrates are highlighted by

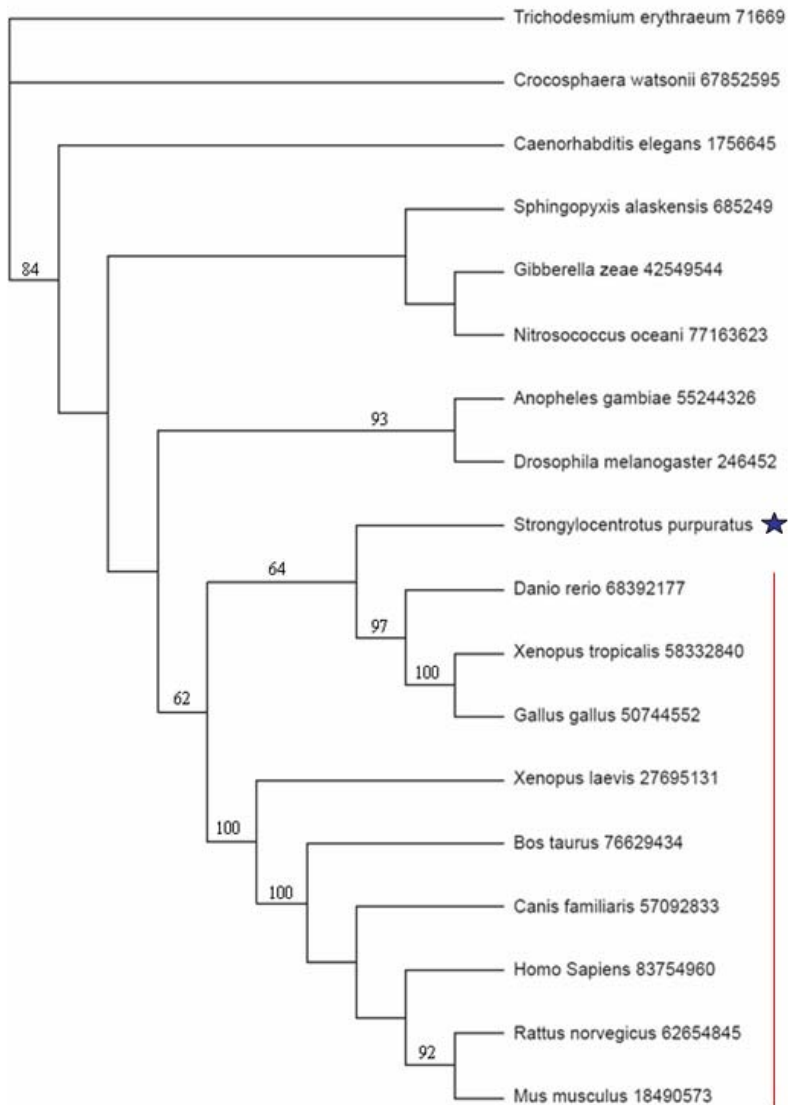the red line and the sea urchin gene is indicated by a blue star.
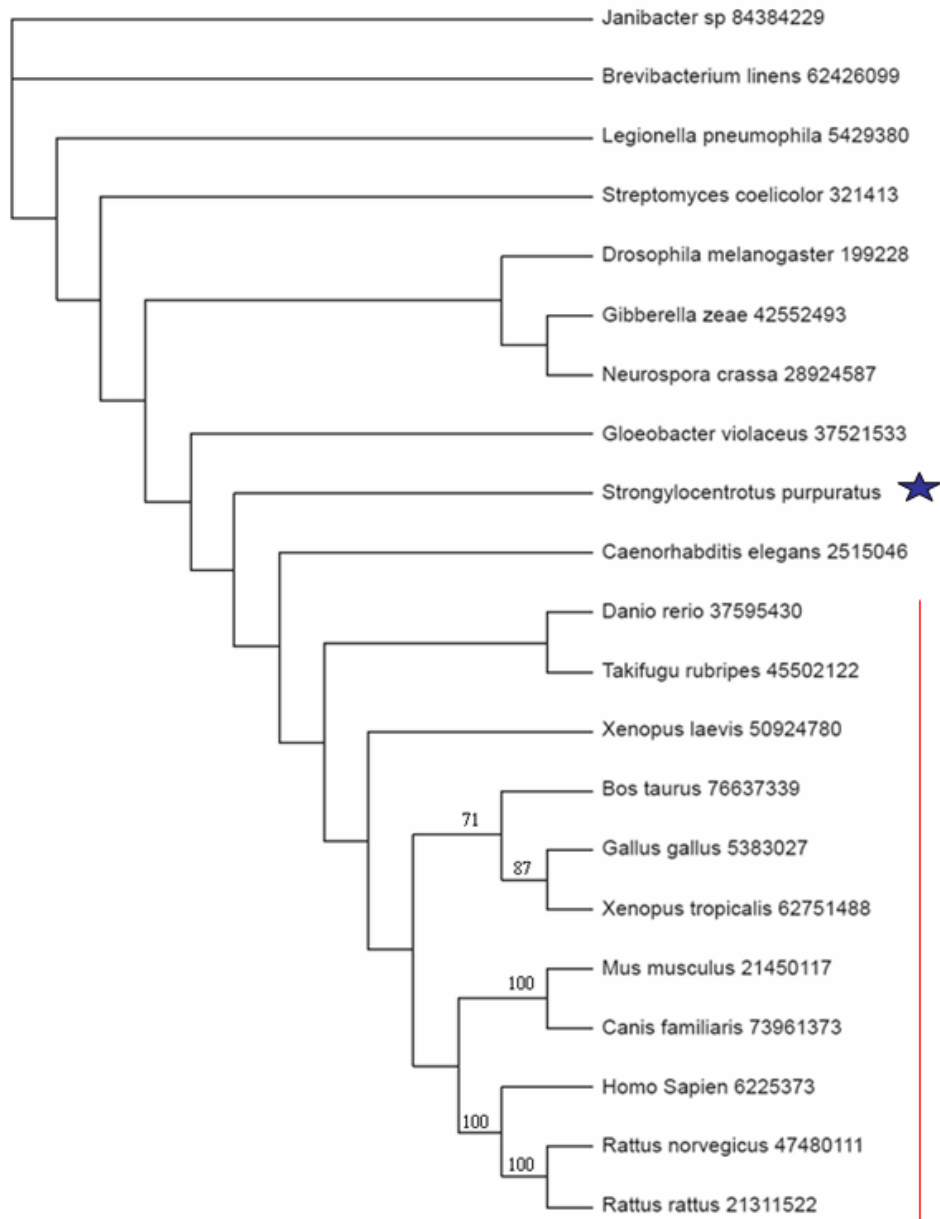
Figure 24: The tree topology of *SpFmo3* was determined through Maximum Parsimony analysis.

The confidence at each node was assessed with 500 bootstrap replications. Vertebrates are highlighted by

the red line and the sea urchin gene is indicated by a blue star.
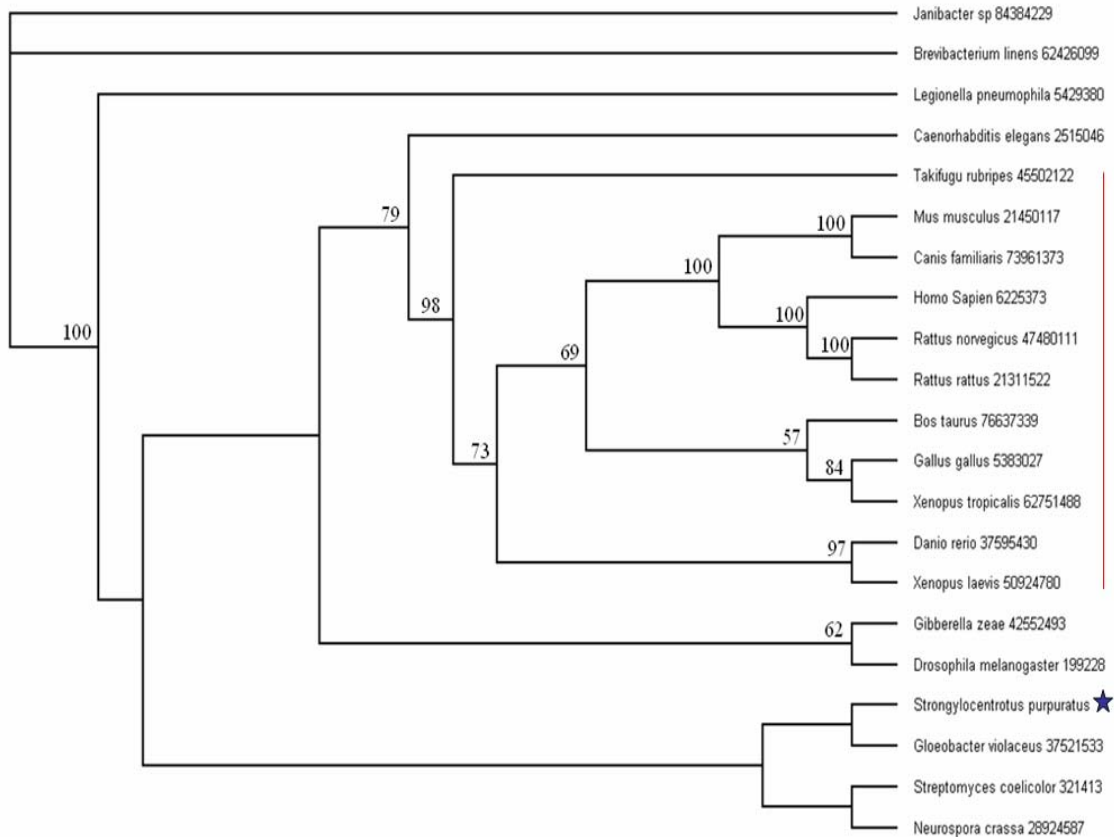
Figure 25: The tree topology of *SpFmo2* was determined through a Neighbor-joining analysis.

The confidence at each node was assessed with 500 bootstrap replications. Vertebrates are highlighted by

the red line and the sea urchin gene is indicated by a blue star.
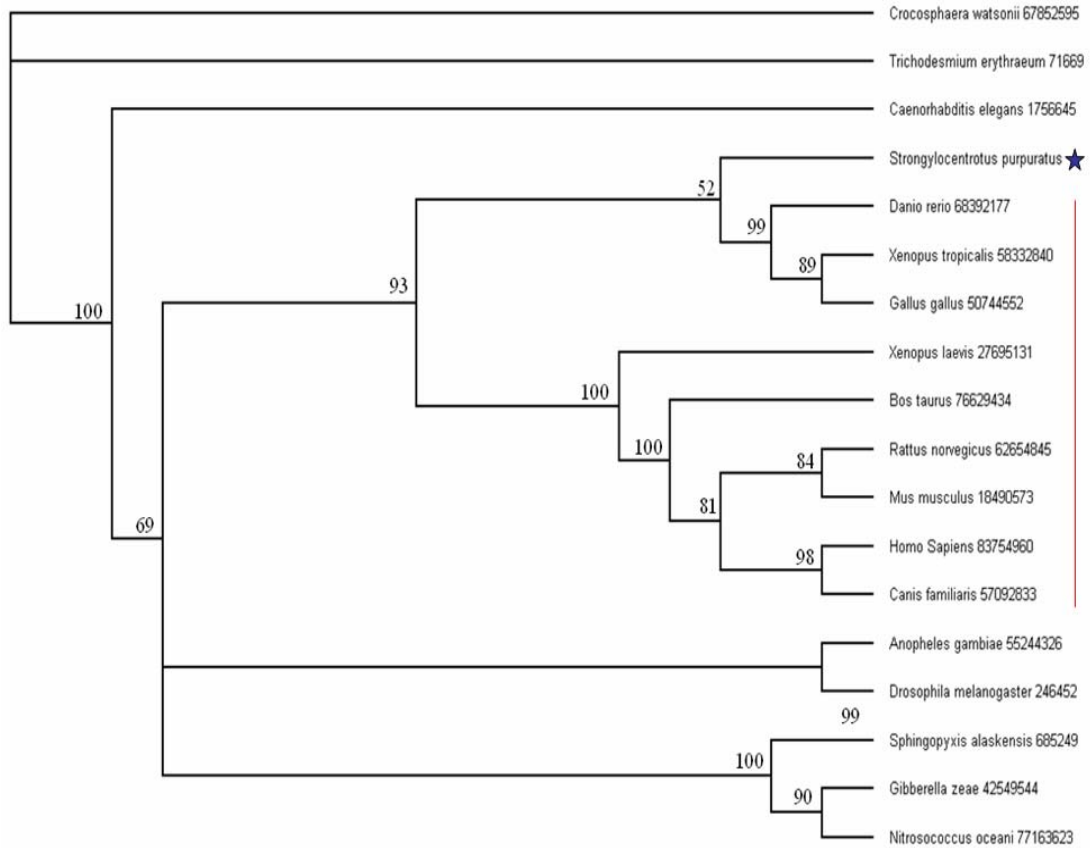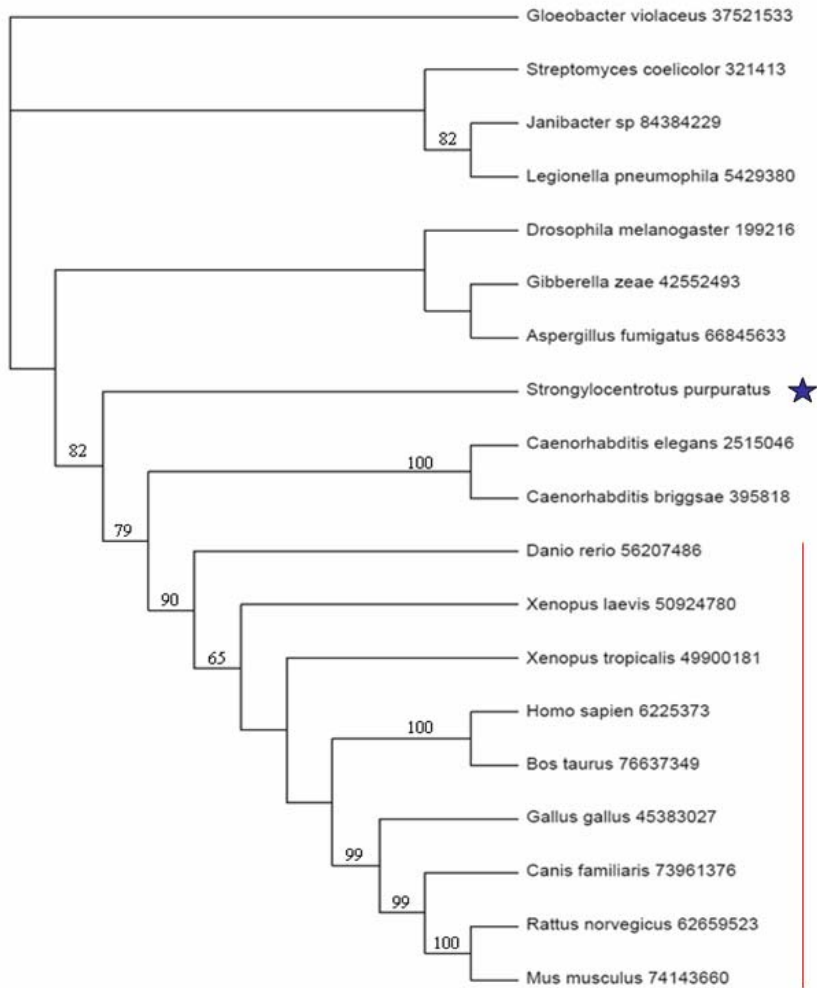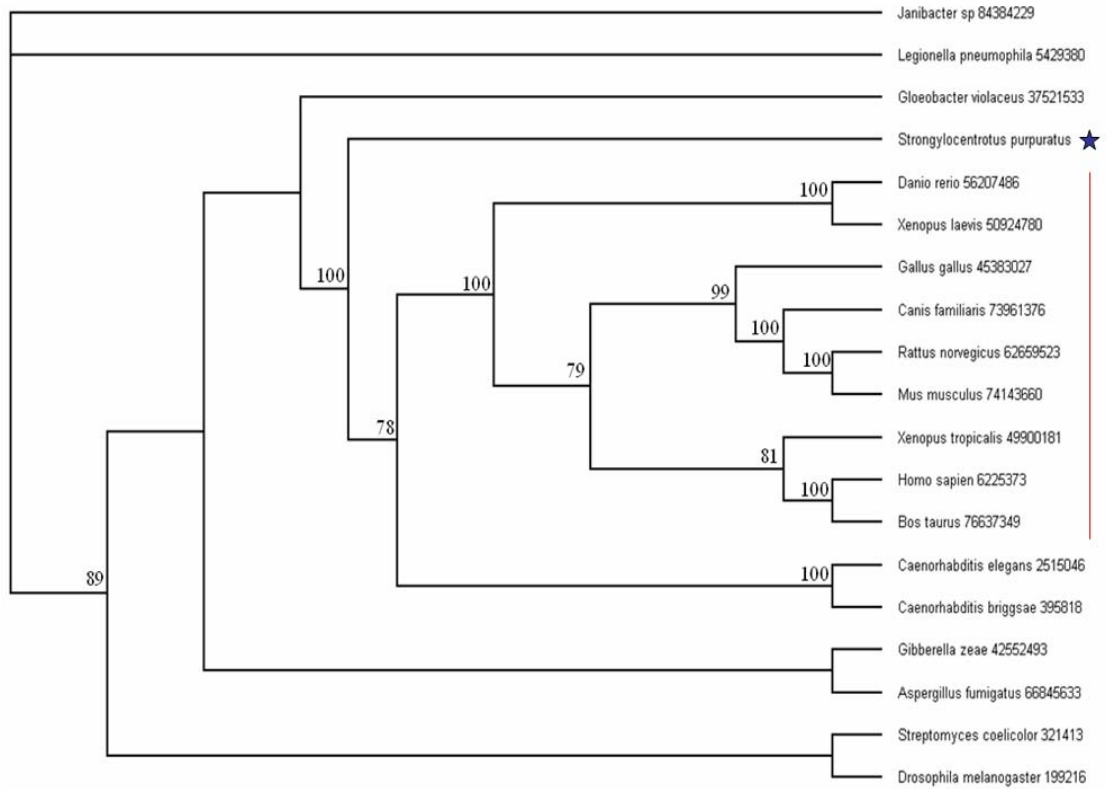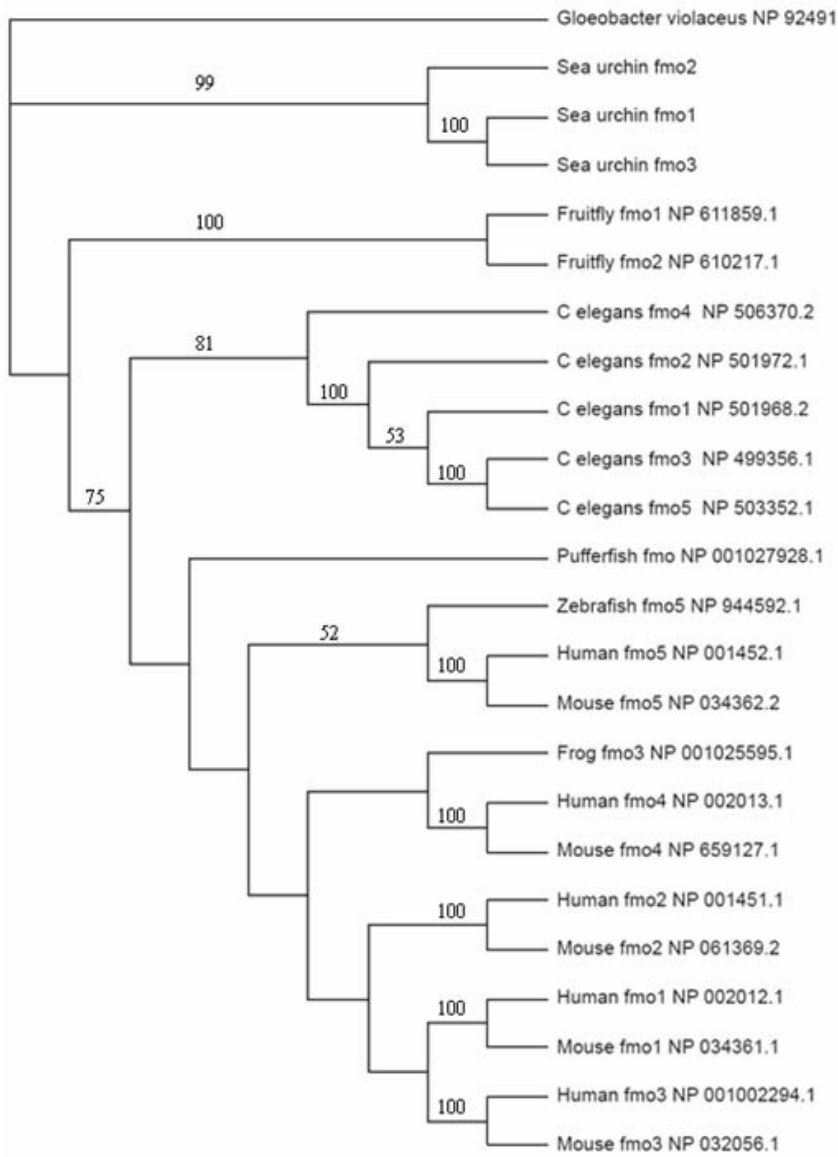
Figure 26: The tree topology for *SpFmo 1-3* was determined through Maximum Parsimony analysis.

The confidence at each node was assessed with 500 bootstrap replications. Vertebrates are highlighted by
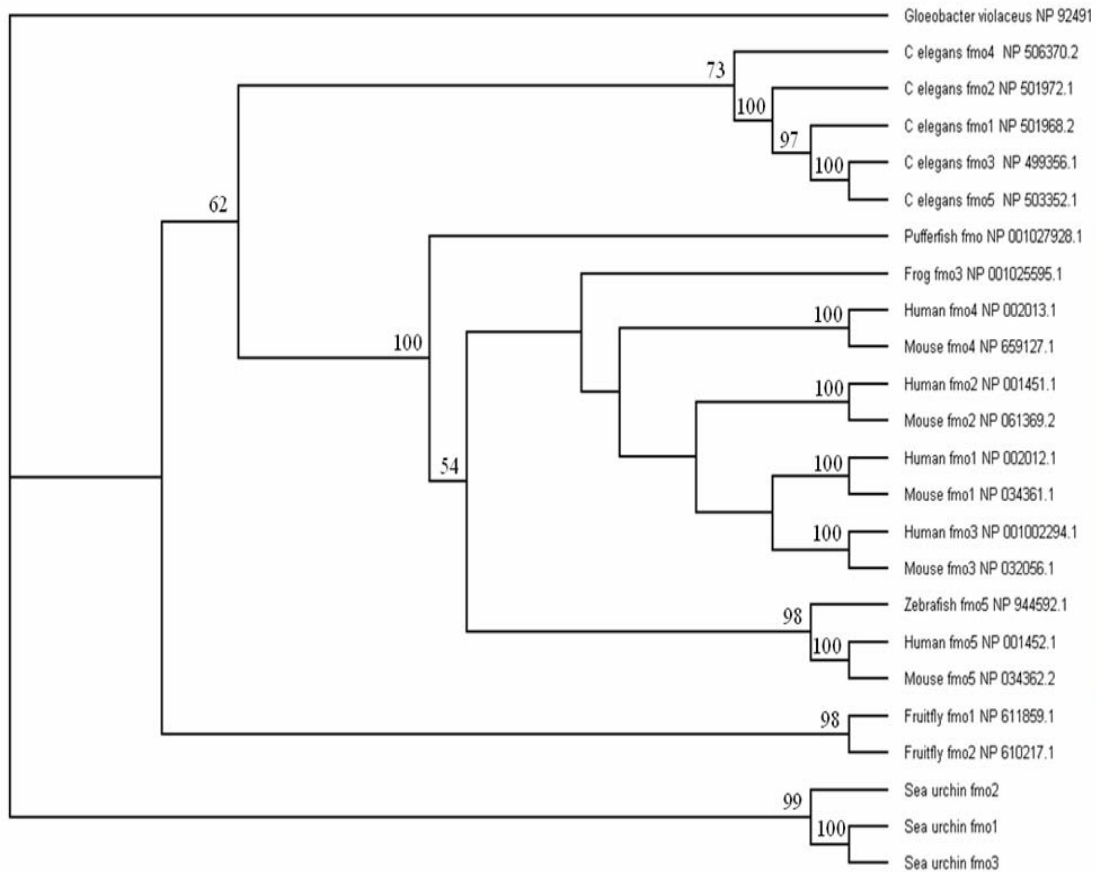
the red line.

Figure 27: The tree topology for *SpFmo1-3* was determined through a Neighbor-joining analysis.

The confidence at each node was assessed with 500 bootstrap replications. Vertebrates are highlighted by

the red line.

Developmental Expression Profile of Pigment Cell Specific Genes

The future direction of this research is to characterize the cis-regulatory regions of these

genes, in particular the co-regulation of the pigment cell differentiation gene battery including

SpPks, *SpSult*, and *SpFmo1-3*. Promoter structure is generally studied by mutagenizing

fragments of the promoter and studying their regulatory properties. The detailed gene expression profile of these genes will serve as a reference for comparing the effects caused by mutagenesis of the promoter.

The transcript prevalence for each gene was measured by Quantitative real-time PCR (QPCR). QPCR is a method that allows quantification of a specific cDNA in a sample. Overall, the methodology is based on a fluorescent reporter molecule, such as Sybr Green, that binds to double-stranded DNA. With each cycle, more PCR product is amplified and the detection of the fluorescence increases. The higher the level of transcript abundance, the faster a significant increase in fluorescence will be detected.

During the first few cycle of QPCR there is very little change in fluorescence detection and these stages identify the baseline. An increased detection of fluorescence above the baseline signifies an increase in PCR product. Data shown are derived from one cDNA batch. The time course was replicated on two other cDNA batches with similar results. The Ct for each primer set was normalized to the ubiquitin Ct for each reaction. Ubiquitin expression is known to be approximately constant throughout development (Nemer et al., 1991 ; Ransick et al., 2002).

SpPks

The onset of transcription for *SpPks* begins at the blastula stage (between 18 and 21 hours) and the level of transcript continues to drastically increase through the start of gastrulation (approximately 30 hours). After 30 hours, the level of transcript decreases by almost 30 times (Figure 28).

Figure 28: Temporal Gene Expression Profile of *SpPks.*

dCt values are along the y-axis and embryonic stages from egg to pluteus are on the x-axis. dCt indicates the difference between the ubiquitin and the gene of interest Ct values. dCt values below -13 indicate no expression of the gene of interest (as indicated by the red line).

SpSult

The gene expression profile for *SpSult* has a trend similar to *SpPks.* The onset of transcription for *SpSult* also begins at the blastula stage (between 15 and 18 hours) and the level of transcript continues to drastically increase to 30 hours. After 30 hours, the level of transcript decreases (Figure 29).

Figure 29: Temporal Gene Expression Profile of *SpSult*.

 dCt values are along the y-axis and embryonic stages from egg to pluteus are on the x-axis. dCt indicates

the difference between the ubiquitin and the gene of interest Ct values. dCt values below -13 indicate no

expression of the gene of interest (as indicated by the red line).

SpFmo1

The onset of transcription for *SpFmo1* begins at the blastula stage (between 12 and 15

hours) and the level of transcript gradually increases through the end of gastrulation

(approximately 40 hours). After 40 hours, the level of transcript decreases (Figure 30).
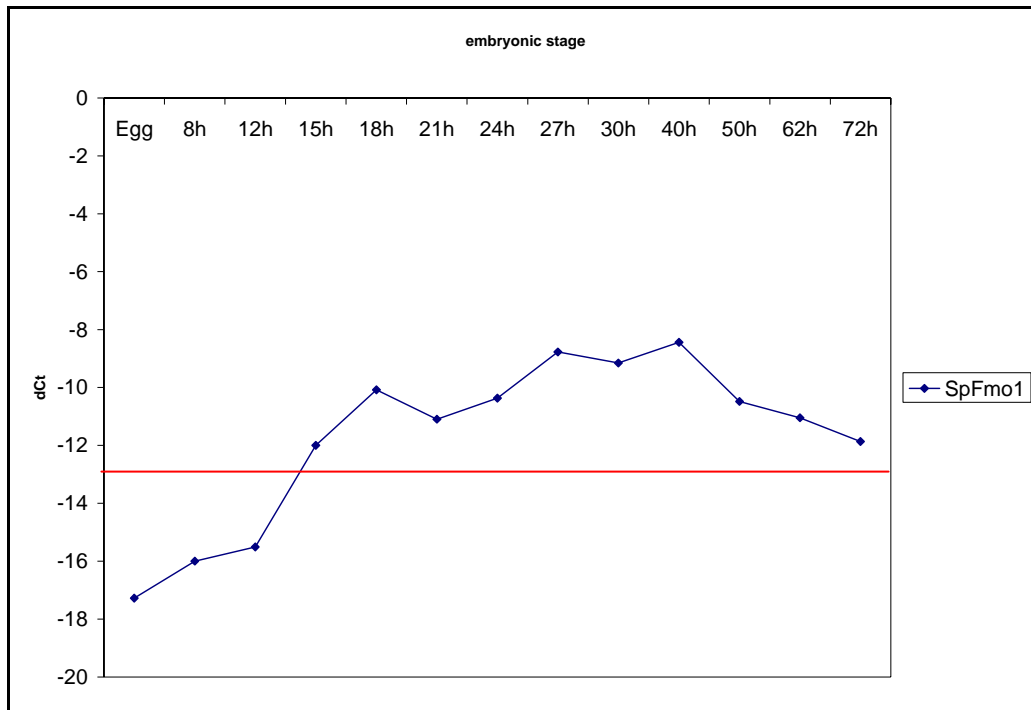
Figure 30: Temporal Gene Expression Profile of *SpFmo1*.

dCt values are along the y-axis and embryonic stages from egg to pluteus are on the x-axis. dCt indicates the difference between the ubiquitin and the gene of interest Ct values. dCt values below -13 indicate no expression of the gene of interest (as indicated by the red line).

SpFmo2

The onset of transcription for *SpFmo2* begins at the blastula stage (between 15 and 18 hours) and the level of transcript continues to increase through 40 hours. After 40 hours, the level of transcript begins to decrease (Figure 31).

SpFmo3

The gene expression profile for *SpFmo3* is very similar to that of *SpFmo1* and *2*. The onset of transcription begins at the blastula stage (approximately 18 hours) and the level of

transcript continues to increase through 40 hours. After 40 hours, the level of transcript begins to slightly decrease (Figure 32).
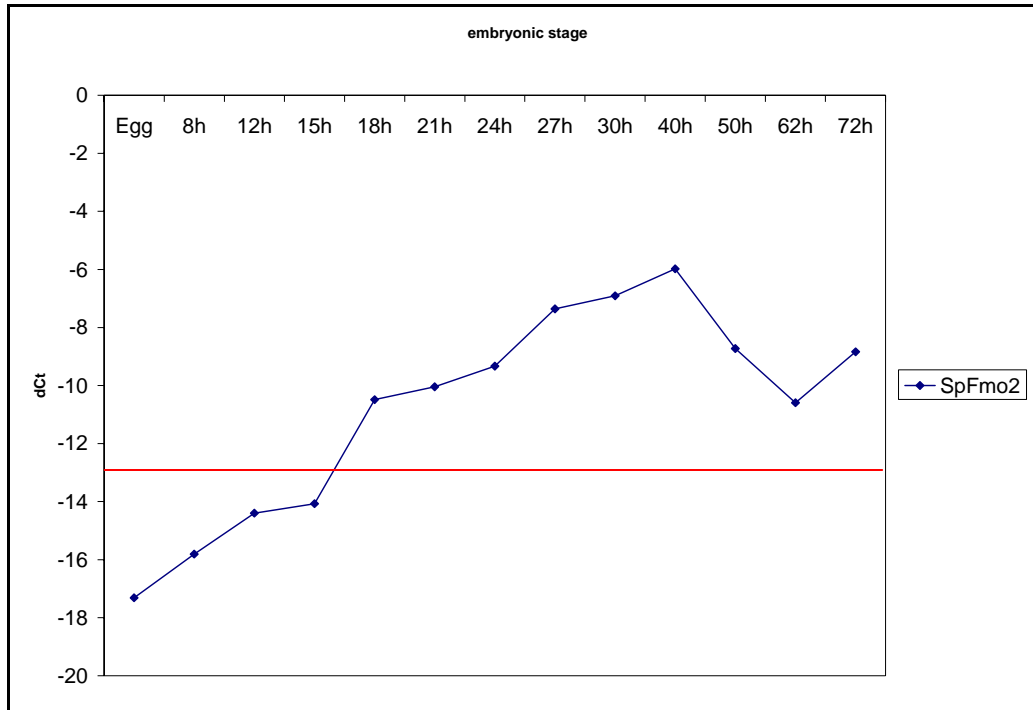
Figure 31: Temporal Gene Expression Profile of *SpFmo2*.

dCt values are along the y-axis and embryonic stages from egg to pluteus are on the x-axis. dCt indicates the difference between the ubiquitin and the gene of interest Ct values. dCt values below -13 indicate no expression of the gene of interest (as indicated by the red line).

SpGcm

The onset of transcription for *SpGcm* begins at approximately 12 hours, about three to six hours before the onset of the differentiation genes described above. The level of transcript

gradually increases until 27 hours where it begins to decrease and by 72 hours no expression is
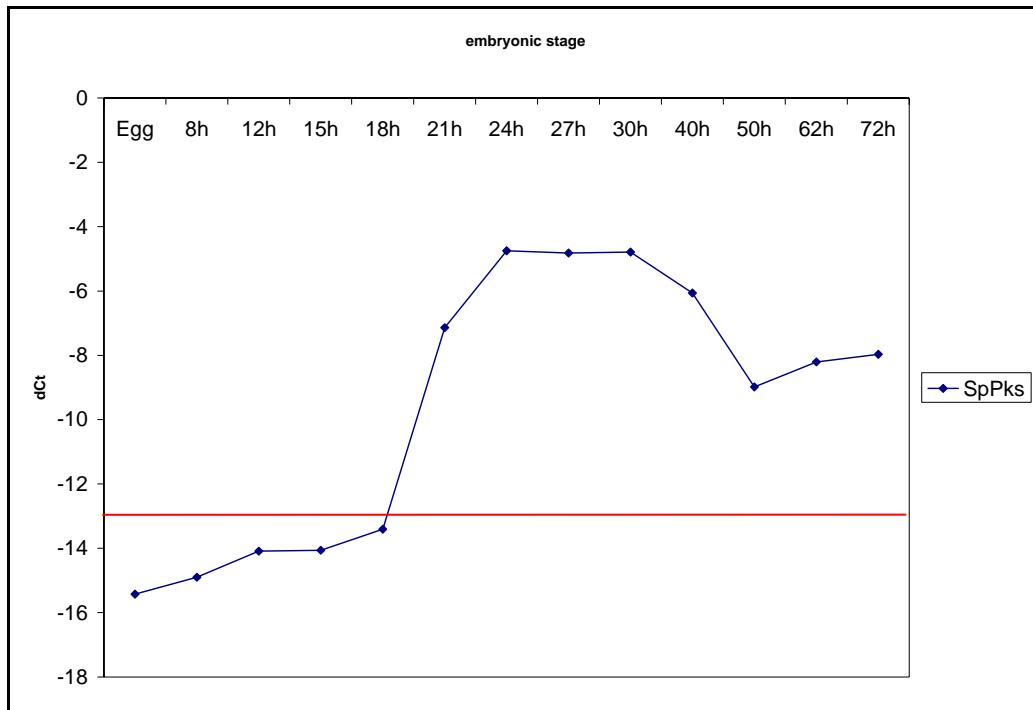
observed (Figure 33).
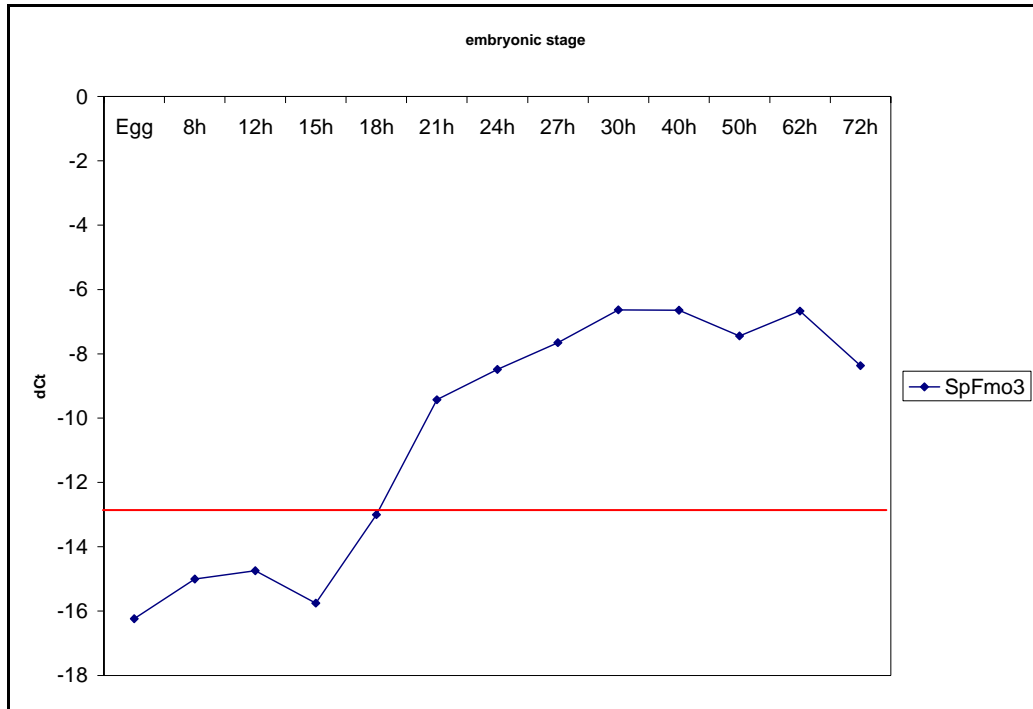
Figure 32: Temporal Gene Expression Profile of *SpFmo3*.

dCt values are along the y-axis and embryonic stages from egg to pluteus are on the x-axis. dCt indicates
the difference between the ubiquitin and the gene of interest Ct values. dCt values below -13 indicate no
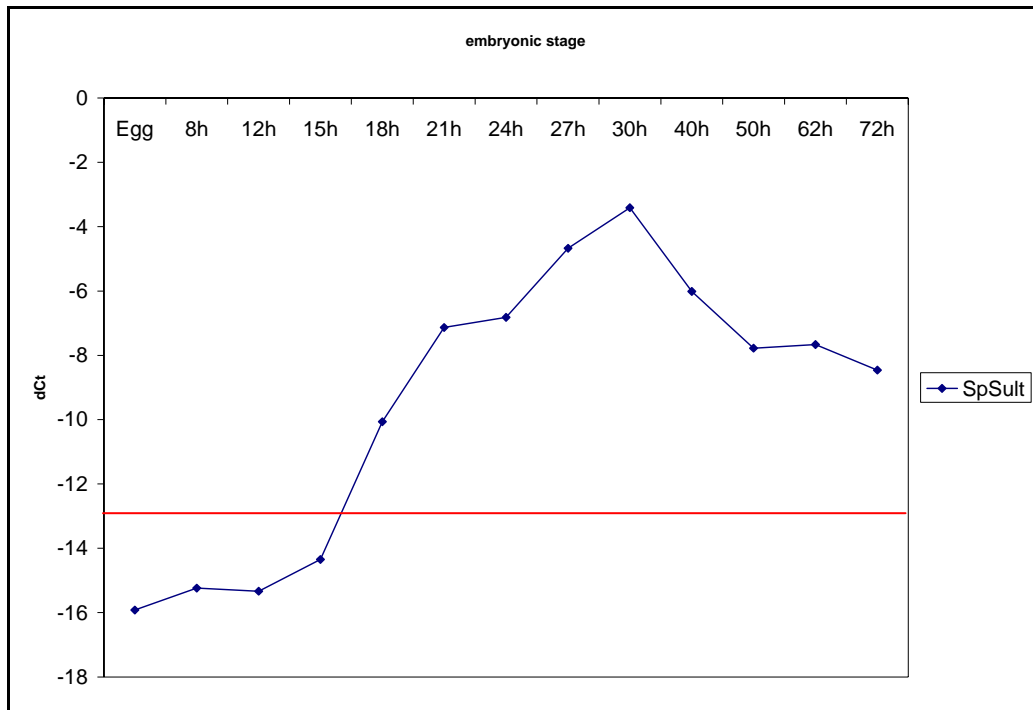expression of the gene of interest (as indicated by the red line).
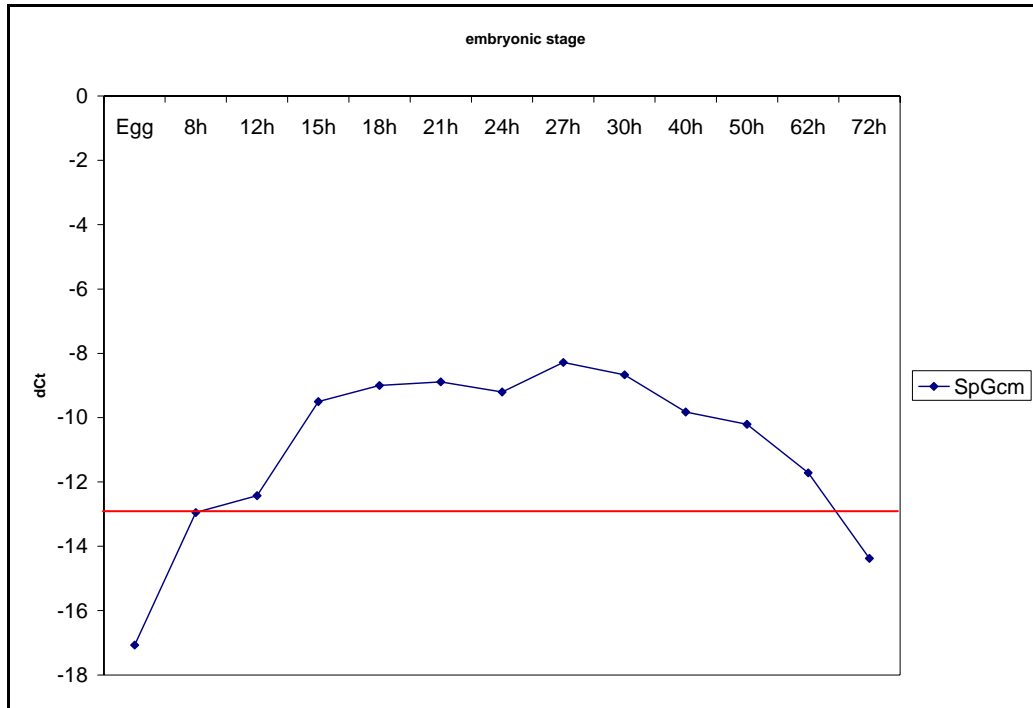
Figure 33: Temporal Gene Expression Profile of *SpGcm.*

dCt values are along the y-axis and embryonic stages from egg to pluteus are on the x-axis. dCt indicates

the difference between the ubiquitin and the gene of interest Ct values. dCt values below -13 indicate no

expression of the gene of interest (as indicated by the red line).

Characterization of Gene Expression of *SpPks*, *SpFmo1*, and *SpFmo3*

In addition to the characterization of the developmental time course by QPCR, a Northern

blot analysis was performed to obtain information about transcript size and about the presence or

absence of alternatively spliced transcripts for each gene. Transcript length determined by the

annotation process was based on partial cDNA sequence and computational exon predictions,

therefore it might not reflect the actual transcript length and it is unable to predict alternatively

spliced isoforms.

The Northern blot analysis was performed at the embryonic stages where the genes were expressed at a relatively high level (based on QPCR data). Data from the Northern Blot Analyses for mRNA transcripts of *SpPks*, *SpFmo1*, and *SpFmo3* at 25 and 40 hours are shown in Figures 34 and 35. The northern blot for *SpPks* showed one 8KB band both at 25 and 40 hours (Figure 34). The northern blot for *SpFmo3* appears to show multiple bands between 1.5 and 2kb (Figure 35).
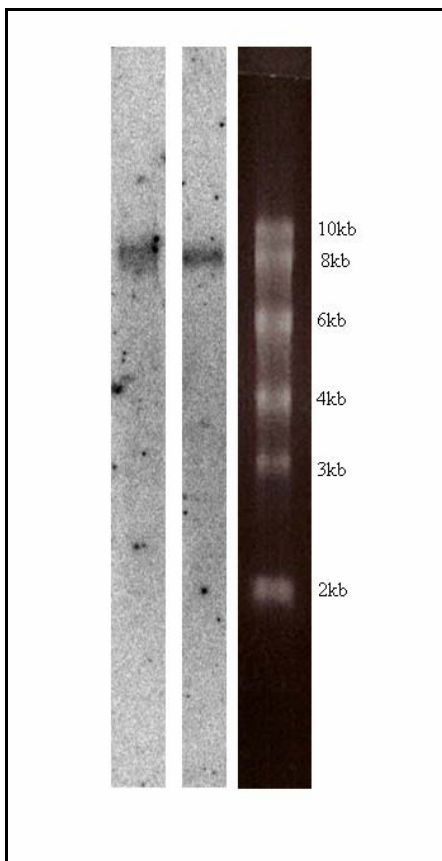
Figure 34: Northern blot analyses shows banding pattern for *SpPks*.

From the left: 25 hours, 40 hours and High Molecular Weight ladder. The transcript size for *SpPks* was extrapolated by aligning the EtBr stained ladder with the autoradiography of the Northern blot.
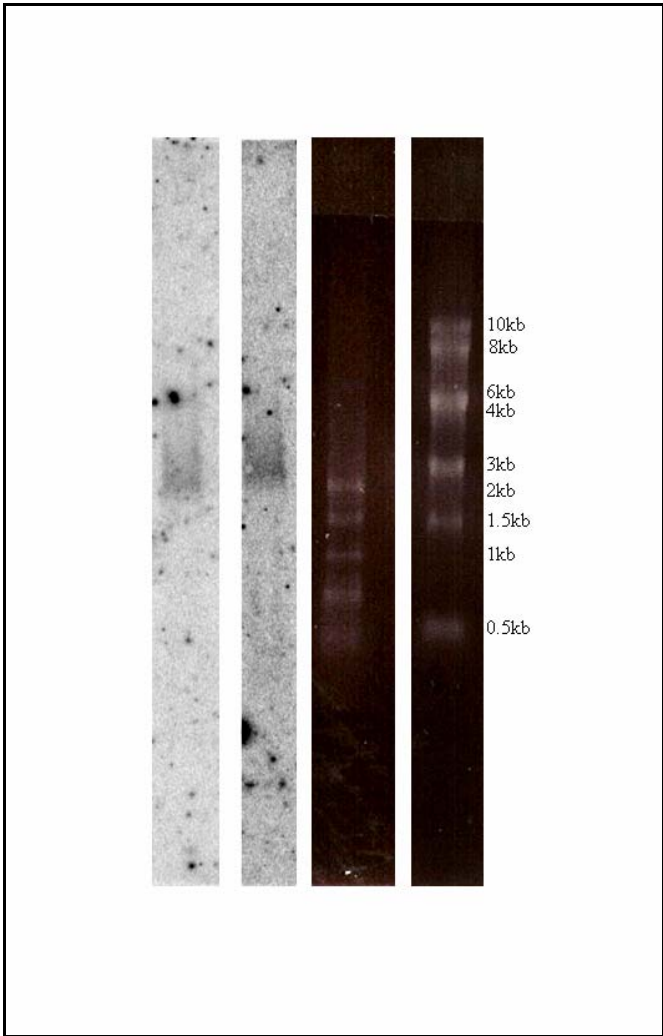
Figure 35: Northern blot analyses shows banding pattern for *SpFmo3*.

From the left: 25 hours, High Molecular Weight ladder, 40 hours and Low Molecular Weight Ladder. The transcript size for *SpFmo3* was extrapolated by aligning the EtBr stained ladder with the autoradiography of the Northern blot.

Temporal and Spatial Analyses using Whole Mount *In Situ* Hybridization

Prior to this work, whole mount *in situ* hybridization (WMISH) was done on the five pigment cell genes at hatched blastula, mid-gastrula, and pluteus stage (Calestani et al., 2003) allowing the visualization of when and in which cells these genes are expressed. In order to create a more detailed temporal and spatial expression analysis of these genes, additional stages were examined by WMISH from blastula throughout the pluteus stage (18, 21, 24, 27, 30, 40, 50, and 63 hours).

SpPks

The staining for *SpPks* was visible from blastula stage (18 hours) [Figure 36]. From blastula (18 hours) prior to gastrulation (27 hours), *SpPks* was expressed in SMC precursor cells at the vegetal plate. At the start of gastrulation (30 hours), the gene was expressed in cells that were beginning to transfer into the blastocoel (pigment cells). During late gastrula, prism (50 hours), and early pluteus stage (63 hours), *SpPks* was detected in cells embedded within the ectoderm.
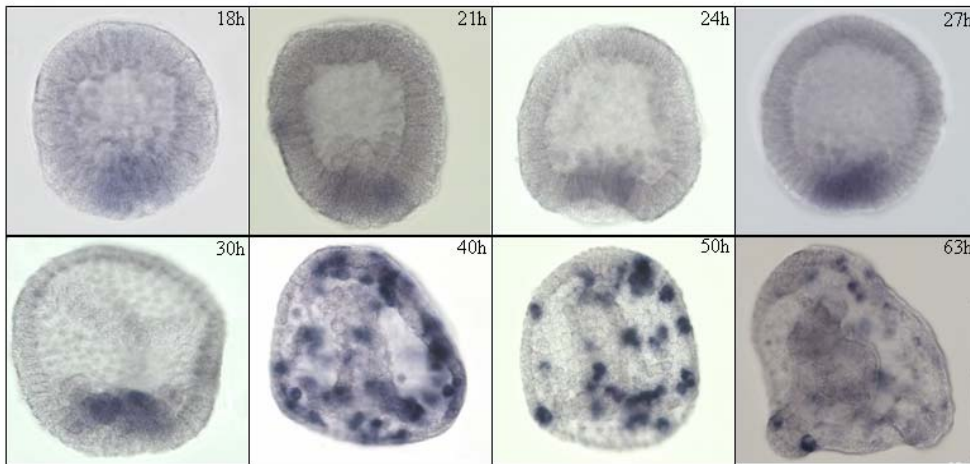
Figure 36: Whole mount *in situ* hybridization of *SpPks*.

Probed with DIG-labeled Antisense RNA from hatched blastula throughout pluteus stage: 18 hours throughout 50 hours are taken at a side view while 63 hours is an oral view.

SpSult, SpFmo1, SpFmo2, and SpFmo3

The staining for *SpSult* and *SpFmo1-3* was only visible in the later stages (40, 50, and 63 hours) [Figures 37 – 40]. It is known that these genes are expressed earlier during blastula stage, from QPCR data and previous WMISH experiments (Calestani et al., 2003). Further experiments will be performed to clarify these results. During late gastrula (40 hours), prism (50 hours) and early pluteus stage (63 hours), the transcripts were detected in cells embedded within the ectoderm.

76
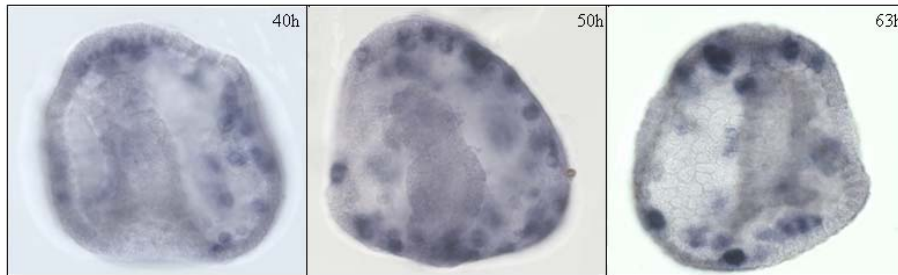
Figure 37: Whole mount *in situ* hybridization of *SpSult*.

Probed with DIG-labeled Antisense RNA from gastrula throughout early pluteus stage: 40 and 50 hours are taken at a side view while 63 hours is an oral view.

Figure 38: Whole mount *in situ* hybridization of *SpFmo1*.

Probed with DIG-labeled Antisense RNA from gastrula throughout early pluteus stage: 40 and 50 hours are taken at a side view while 63 hours is an oral view.
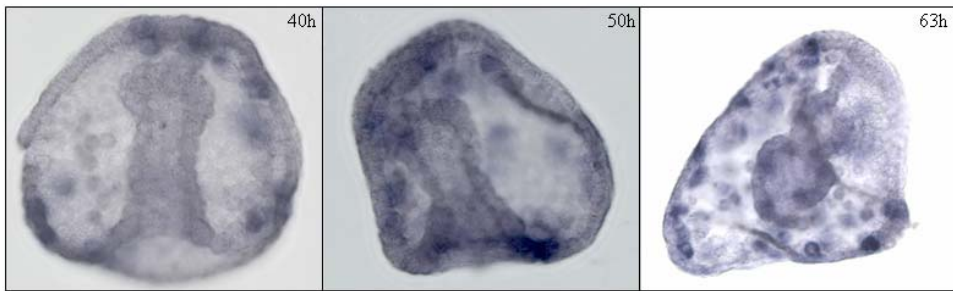
Figure 39: Whole mount in situ hybridization of *SpFmo2*.

Probed with DIG-labeled Antisense RNA from gastrula throughout early pluteus stage: 40 and 50 hours are taken at a side view while 63 hours is an oral view.

Figure 40: Whole mount *in situ* hybridization of *SpFmo3*.

Probed with DIG-labeled Antisense RNA from gastrula throughout early pluteus stage: 40 and 50 hours are taken at a side view while 63 hours is an oral view.

Functional Study of *SpFmo1* and *SpSult* by Microinjection of Morpholino Antisense Oligonucleotides

Functional analyses using Morpholino antisense oligonucleotides (MASOs) is one of the most informative approaches for reverse genetics by knocking down a gene. The MASOs cause a loss of function accomplished by preventing cells from producing a specific protein. MASOs

bind to the leader sequence and first 25 bases of a targeted RNA sequence. Once bound to the 5' UTR region, MASOs sterically block the translation initiation complex from the 5' cap to the start codon thus inhibiting the translation of a specific transcript. Prior to this study, antisense technology showed that *SpPks* and *SpFmo1-2* were essential for the biosynthesis of echinochrome (Calestani et al., 2003). The knockdown of these genes resulted in embryos that did not display any pigment. In order to understand any potential roles of *SpFmo3* and *SpSult* in pigment biosynthesis, functional analyses were performed using MASOs to block these genes. The results failed to show any morphological abnormalities in the knockdown embryos.

DISCUSSION

In this work a group of mesodermal sea urchin genes, *SpPks, SpSult*, and *SpFmo1-3*, which are expressed during the specification and differentiation processes of pigment cells (*SpPks*, *SpFmo* and *SpSult*; Calestani et al., 2003) were characterized in terms of gene structure, gene expression pattern during embryo development and phylogenetic history. These data provide the ground work for future studies of the Gene Regulatory Network (GRN) underlying pigment cells development. In the recent years, the sea urchin embryo has become a robust model system for understanding transcriptional regulation during development. At present, the sea urchin GRN for endoderm and mesoderm specification (Reviewed in Davidson et al., 2002) and the dorsal-ventral patterning in *Drosophila* (reviewed in Levine and Davidson, 2005) are the two most extensively characterized developmental GRNs.

Structural and functional gene annotation of *SpPks*, *SpFmo1-3* and *SpSult*

Polyketide synthases (PKSs) are mainly found in bacteria, fungi, and plants. They are typically responsible for the biosynthesis of a wide variety of polyketide compounds (reviewed in Hopwood, 1997; Hopwood and Sherman, 1990; Stuanton and Weissman, 2001). In the sea urchin it is known that *SpPks* is required for the biosynthesis of the pigment echinochrome, a compound that has conveyed antiobiotic properties against bacteria (Calestani et al., 2003). In sea urchin embryos, *SpPks* is expressed in pigment cells beginning at the hatched blastula stage and it is maintained throughout the pluteus stage (Calestani et al., 2003).

The transcript for *SpPks* is quite large, 8, 447 base pairs in length. Large open reading frames (ORFs) are typical of multifunctional PKSs. The comparison of SpPKS sequence with the sequences deposited in the NCBI Conserved Domain Database led to the identification of a

polyketide synthase domain (KS), an acyltransferase domain (AT), and a zinc-binding

dehydrogenase domain. The first domain, the polyketide synthase domain, typically polymerizes

acetyl-CoA or malonyl-CoA. This is done by succeeding decarboxylating Claisen condensations.

Following the pks domain is the acyltransferase domain. Acyltransferases catalyze the transfer of

acyl groups from donor to acceptor, forming either esters or amides. The third domain has

sequence similarity to a zinc-binding dehydrogenase domain. A general classification commonly

used for this type of PKS domain is enoyl reductase (ER). Zinc-containing alcohol

dehydrogenases bind atoms of zinc and catalyze the reversible oxidation of ethanol to

acetaldehyde with the concomitant reduction of NAD. Based on what is known about the

biochemistry of PKSs, the presence of a reduction step requires a dehydration step, which is

usually performed by a dehydratase domain (DH). This domain is not very conserved

(HXXXXXXXXP), therefore the blast comparison in the NCBI Conserved Domain Database did

not detect it, but a DH domain was indeed found between the AT and ER domain while

analyzing the SpPKS amino acid sequence. Another domain present at the C-terminus of the

protein is a phosphopantetheine attachment site (PP-binding) domain which is very short and not

very conserved (GXXS). The PP-binding domain is a component of acyl carrier proteins where it

functions as a "swinging arm" for attaching acyl groups and amino acid groups (Pugh and Wakil,

1965). The PP-binding domain is present in PKS types I and II but not in type III. *SpPks* encodes

the following series of conserved domains: KS-AT-DH-ER-PP (Calestani et al., 2003). The

SpPKS domains are part of one multifunctional polypeptide (according to the cDNA sequence),

therefore it is concluded that SpPKS is type I. In contrast, type II PKSs are composed of

numerous monofunctional subunits in which the active sites are dispersed amid several

polypeptides (Hopwood, 1997; Shen, 2000; Staunton and Weissman 2001; Khosla et al. 1999;

Keating and Walsh 1999; Hutchinson 1999). Type III PKSs can also be ruled out because they do not contain a PP-binding domain (Moore and Hopke, 2001; Gross et al., 2006).

Sulfotransferases are found in bacteria through higher eukaryotes and are responsible for catalyzing the sulfate conjugation of several substrates. The sulfate conjugation results either in compound detoxification or bioactivation. In mammals, sulfate conjugation is important in the detoxification of drugs and xenobiotics (Falany, 1997; Weinshilboum et al., 1997). In plants, water soluble plant pigments called flavonols, as well as other endogenous compounds, undergo sulfate conjugation (Varin and lbrahim, 1989). Typically, the resulting sulfonation increases the water solubility of these compounds, thus allowing excretion. In mammals, sulfonation is also an important mechanism to regulate the activity hormones and neurotransmitters (Falany, 1997; Weinshilboum et al., 1997). In these instances, bioactivation occurs because the sulfate group generated by sulfonation is electron-withdrawing and capable of binding to other cellular macromolecules (Minchin et al., 1992; Chou et al., 1995)

Sulfotransferases illustrate different developmental and tissue-specific gene expression (Her at al., 1997; Dunn and Klaassen, 1998). For example, in humans, there are three families of sulfotransferases composed of thirteen members: SULT1 - A1, A2, A3, A4, B1, C2, C4, E1; SULT2 - A1 and B1; SULT4A1 (Blanchard et al., 2004; Hildebrandt et al., 2004). In the SULT1 family, *sult1A3* is expressed in the liver, kidney, lung, small intestine, spleen and leukocytes while *sult2A1*, a member of the SULT2 family, is expressed solely in the liver and adrenal glands. The sea urchin sulfotransferase gene is expressed solely in pigment cells, suggesting that it could play a specific role in pigment cell specific biosynthetic processes. The transcript for *SpSult* is 858 base pairs in length. Although seemingly small, the putative orthologues retrieved from the NBCI database had similar transcript lengths.

FMOs are mainly found in bacteria, plants, and higher metazoan. These NADPH-dependent flavoproteins are responsible for catalyzing the oxidation of several compounds including nucloephilic nitrogen, sulphur, phosphorous and selenium atoms. FMOs are primarily involved in the detoxification of various xenobiotics, as well as the molecular activation of numerous metabolites. Three different sea urchin *fmo* genes have been isolated that may be involved in echinochrome biosynthesis (Calestani et al., 2003). In mammals, five *fmo* have been isolated thus far (Hines at al., 1994; Gasser, 1996; Cashman, 2000).

Mammalian FMOs are highly variable. They display different developmental, tissue-specific, and substrate specificities (Dolphin et al., 1996; Dolphin et al., 1998; Koukouritaki et al., 2002). In humans, *fmo1* is found in the adult kidney and fetal liver, *fmo2* is expressed mainly in the adult lung, *fmo3* is expressed in the adult liver, and *fmo4-5* is found in the fetal and adult liver. In tissues where FMO is copious (i.e. adult kidney and liver), the activity is typically associated with immunoreactivity and is controlled by the most prominent FMO form (Lattard et al., 2004). In tissues where FMO expression is low (i.e. brain), the FMO function is unclear and the characterization remains a challenge (Lattard et al., 2004).

SpFmo1, *SpFmo2* and *SpFmo3* transcripts have similar lengths, respectively 1,949, 1,192 and 2,200 base pairs. It may be possible that *SpFmo1*, *SpFmo2*, and *SpFmo3* are involved in various catalytic steps of echinochrome biosynthesis. Considering the relatively high level of oxygenation of echinochrome, SpFMOs could be involved in oxygenating further the precursor synthesized by SpPKS. The three sea urchin *fmo* genes are expressed in the same cell type (pigment cells) and their expression begins at the same time (at about 15 hours). Together these results suggest that the three *SpFmo* genes may be co-regulated by a commen set of transcription factors.

Phylogenetic Analysis of *SpSult* and *SpFmo1-3*

Due to its phylogenetic placement between vertebrates and invertebrates, the sea urchin fills in a significant evolutionary gap between the deuterostome chordates and non-deuterstomes. This fulfillment has allowed further insight into the evolution that occurred subsequent to the split between the ancestors that gave rise to humans and insects. Phylogenetic analyses were conducted on *SpSult* and *SpFmo1-3* using both character and distance based approaches. By making use of two different approaches, it was possible to cross-validate phylogenetic estimations since these two approaches have different advancements and assumptions about evolutionary processes.

Initial BLAST results of the sea urchin SpPKS against the NCBI database revealed low sequence similarity to FASs in human, *Drosophila*, and *Caenorhabditis* and failed to reveal any other obvious orthologues in other animal genomes. The initial comparisons did show, however, high sequence similarity with fungal and bacterial sequences. The phylogenetic analyses performed included several eubacterial and fungal sequences, *Pks* like genes in the chicken, fish, fruit fly, mosquito, and *Fas* like genes in the chicken, fish, human, dog, mouse, rat, and nematode. The analyses suggested that *SpPks* is most closely related to a *Pks* found in the slime mold *Dictyostelium discoidium* (Castoe et al., 2006).

D. discoidium typically live on the forest floor and feed off of bacteria. When their source of food becomes sparse, *D. discoidium* amebas join together and transform into a multicellular tower which moves to a warmer environment. The transformation is triggered by Differentiation Inducing Factor (DIF-1). Once signaled, cells metamorphose into next-generation spore cells or stalk cells and wait for the food source to flourish. The chemical structure of DIF-1 is similar to natural plant products (chalcones). DIF-1 is biosynthesized by a hybrid type I-type III PKS

(Austin et al., 2006). The type I region of *D. discoidium* PKS is the one having high similarity with the SpPKS. DIF-1 and the sea urchin echinochrome are both relatively small polyketide compounds but they are chemically different in terms of number of carbon units, 15 and 12 respectively, and level of reduction. Therefore they are expected to have relatively different biological function. Phylogenetic analyses conducted on *SpPks* suggested a complex evolutionary history of this gene involving either massive gene loss (most likely) or Horizontal Gene Transfer (HGT), a rare event in which an organism transfers genetic material to another cell that is not its offspring (Castoe et al., 2007). Phylogenetic analyses were conducted on *SpSult* and *SpFmo1-3* in order to evaluate the relationships of these other sea urchin pigment cell genes, belonging to the differentiation gene battery including SpPks, to putative orthologues from other taxa.

For *SpSult*, both trees show the gene as being most closely related to vertebrates, consistent with the tree of life suggesting vertical transfer, when an organism inherits genetic material from an ancestor from which it has evolved (Figures 18 and 19). For *SpFmo1*, both trees suggest that the sea urchin is most closely related to the nematode and vertebrates (Figures 20 and 21). Sea urchins are echinoderms, the closest known relative of chordates. Due to this, it would be more logical for *SpFmo1* to be more closely related to vertebrates while the nematode is more closely related to other protostomes (i.e. *Drosophila*). However, it may be possible that the nematode gene was under strict environmental selective pressure resulting in the gene being more similar to vertebrates than to other protostomes (i.e. *Drosophila*). However, convergent evolution, might be the explanation to the similarity among the nematode, vertebrate and echinoderm *Fmo* genes.

For *SpFmo2*, the Maximum Parsimony tree suggests that the sea urchin gene is most closely related to fungi suggesting the possibility of HGT (Figure 22). Previous studies have shown that HGT has been a significant evolutionary mechanism within prokaryotes and unicellular eukaryotes (Jain et al., 1999; Bapteste et. al, 2005). Currently it is uncertain how relevant HGT is in higher plants and animals (Richardson and Palmer, 2007). Due to the rareness of HGT in multicellular eukarotes, the MP tree positioning of *SpFmo2* is not likely due to HGT. This position is also weakly supported by bootstrap values (<50%). A value greater than 70% indicates significant support for a particular clade (Soltis and Soltis, 2003). Alternately, the Neighbor-joining tree suggests that *SpFmo2* is most closely related to the nematode and vertebrates, possibly due to the misplacement of the nematode or convergent evolution (Figure 23).

The Maximum Parsimony tree for *SpFmo3* also suggests that the sea urchin gene is most closely related to the nematode and vertebrates while the Neighbor-joining tree suggests it is more closely related to bacteria (Figures 24 and 25). Once again, this could suggest horizontal gene transfer however that is unlikely due to the rareness of the event and lack of support for this placement.

A gene family tree for *SpFmo1-3* was constructed in order to see if the sea urchin genes were orthologous to other species. Gene family orthologues have evolved through a speciation event and as a result are more likely to conserve protein function. Paralogous family members, on the other hand, have evolved through gene duplication events and typically do not have similar function (Reviewed in Thornton and DeSalle, 2000; Zmasek and Eddy, 2001). Both gene family trees were consistent in suggesting that the three sea urchin *Fmo* genes are most closely

related to each other than to other species and likely evolved from two successive gene duplication events (Figures 26 and 27).

Characterization of the pattern of gene expression of *SpPks*, *SpFmo1-3* and *SpSult* during sea urchin embryo development

The onset of transcription for the five pigment cell specific genes occurs at the blastula stage (between 15 and 18 hours), a few hours after the activation of *SpGcm* (12 hours). It is known that *SpGcm* is required for the activation of transcription for *SpPks*, *SpSult*, and *SpFmo1-3* (Davidson et al., 2002). Overall, the expression profile for *SpGcm* and the five pigment cell specific differentiation genes suggest that *SpPks*, *SpSult*, and *SpFmo1-3* could be directly activated by *SpGcm*. The few hour difference between the onset of transcription for *SpGcm* and the five differentiation genes suggests the possibility that those five genes may be members of a differentiation gene battery that is co-regulated by the transcription factor SpGCM (Figure 41). *SpPks* and *SpSult* appear to have similar profiles characterized by a drastic increase of transcript abundance between 15 and 24 hours followed by a decrease in the level of transcript prior to the start of gastrulation (approximately 30 hours) [Figures 28 and 29]. This decrease of transcript around 30 hours is concurrent with the decrease of transcript abundance for *SpGcm* and suggests that the pigment cell genes are down-regulated by *SpGcm*. At this stage of development, pigment cells begin to migrate into the blastocoel. This morphogenetic event lead pigment cells to detach from their original epithelium and therefore they are no longer eceiving cell signaling from the surrounding endoderm. This might contribute to the lower level of transcript accumulation.

Likewise, *SpFmo1-3* also show similar profiles characterized by a more gradual accumulation of transcript abundance compared to *SpPks* and *SpSult* (Figures 30, 31, and 32).

From blastula throughout gastrulation (approximately 40 hours), transcript abundance gradually increases followed by a decrease in the level of transcript after 40 hours. This also shows a correlation between the down-regulation of *SpGcm* and the down-regulation of pigment cell genes, suggesting that *SpGcm* is required not only for the onset of transcription but also for the maintenance of these pigment cells' gene expression.
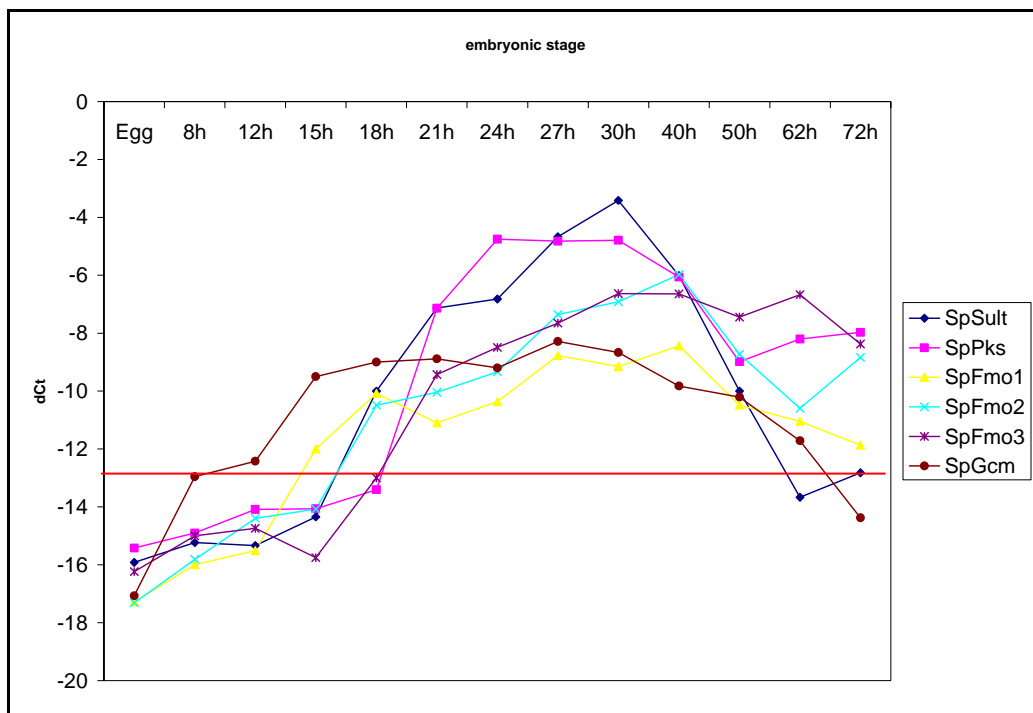
Figure 41: A combined Temporal Gene Expression Profile.

dCt values are along the y-axis and embryonic stages from egg to pluteus are on the x-axis. dCt indicates the difference between the ubiquitin and the gene of interest Ct values. dCt values below -13 indicate no expression of the gene of interest (as indicated by the red line).

In addition, given that in *S. purpuratus* embryos, echinochrome synthesis occurs at the early gastrula stage (Griffiths, 1965) it may be possible that enough protein is already produced

by 30-40 hours, resulting in a faster mRNA turnover, which decreases the transcript levels. It would be interesting to compare the expression profiles of these genes with PMC specific genes (i.e. SM50 and SM60), since PMCs also migrate into the blastocoel (around 24 hours post-fertilization) and it would also be interesting to see if in this case cell migration correlates with drastic changes in gene expression.

In addition to the characterization of the developmental time course by QPCR, a Northern blot analysis was performed to obtain information about transcript size and about the presence or absence of alternatively spliced transcripts for each gene. A Northern Blot Analysis was performed at the embryonic stages where the genes were expressed at a relatively high level (based on QPCR data). Data from the Northern Blot Analyses for mRNA transcripts of *SpPks* and *SpFmo3* are shown in Figures 34 and 35. The northern blot for *SpPks* showed a band around 8kb which is consistent with the size of the mRNA transcript (8,447 base pairs) derived from the gene annotation. The northern blot for *SpFmo3* showed multiple bands between 1.5 and 2kb. This banding pattern suggests the possibility of alternative splicing or multiple alternative 5' ends.

During alternative splicing, exons are either preserved in the mature mRNA or excised in different arrangements that create several mRNAs derived from a single pre-mRNA (Reviewed in Lopez, 1998). Since 2000, several experiments have demonstrated that it is possible to monitor alternative splicing using macroarrays and probes situated at exon-exon boundaries (Yeakley et al., 2002; Hu et al., 2001, Shoemaker et al., 2001; Clark and Sugnet et al., 2002; Wang et al., 2003, Castle et al., 2003; Black, 2000). Currently, there is a macroarray cDNA library that contains all genes expressed at 20 hours post fertilization in the sea urchin. For *Fmo3*, probes could be used to monitor the exon-exon junctions. Then, splicing predictions

derived from the array data could be used to direct reverse transcriptase PCR (RT-PCR) and sequencing validation to specific transcription locations (Johnson et al., 2003).

Multiple 5' ends is a result of multiple transcription start sites in a region of 20-200 base pairs and gives rise to mRNAs with multiple 5' ends. In order to test for this it would be necessary to perform Rapid Amplification of cDNA ends (RACE). This particular PCR-based method was created to assist in the cloning of full length 5' and 3' ends from partial cDNA sequences (Schaefer, 1995). By performing a randomly-primed reverse transcription reaction, an adaptor would be added to the 3' end of the cDNA synthesized (5' end of gene sequence) by ligation or polymerase extension. The amplification by PCR would incorporate a gene specific primer with another primer that recognizes the adapter sequence. When the reaction is completed gel electrophoresis would indicate the presence of one or more transcripts.

Temporal and spatial expression analyses were done on *SpPks*, *SpSult*, and *SpFmo1-3* by whole mount *in situ* hybridization based on the Harkey method and optimized for sea urchin embryos (Harkey, 1992; Ransick et al., 2002). The stages examined were from blastula throughout the pluteus stage (18, 21, 24, 27, 30, 40, 50, and 63 hours).

The gene expression patterns at all stages was consistent with what is already known (temporally and spatially) about pigment cell precursors and differentiated cells distribution within the embryo (fate map produced by Ruffins and Ettensohn, 1996). The staining for *SpSult* and *SpFmo1-3* was only visible in the later stages (40, 50, and 63 hours) in cells embedded within the ectoderm (Figures 37 – 40). Although staining visibility only become apparent after gastrulation, QPCR data and previous WMISH data on 20 hours embryos (Calestani et al., 2003) show that these genes are expressed beginning at the blastula stage. For some genes, especially when expressed at low levels, it is difficult to achieve strong signals while maintaining low

90

background. In particular, early cleavage stages tend to be more prone to higher background staining using this method. Alternatively, a whole mount *in situ* will be repeated on the stages prior to the end of gastrulation (< 40 hours) using a different method to provide a better background signal resolution (Arenas-Mena et al., 2000).

Functional Analysis of *SpSult* and *SpFmo3*

Previous functional analyses of *SpPks* and *SpFmo1*, as well as preliminary data on *Fmo2*, reveal that they are involved in the metabolic pathway of larval pigment biosynthesis. The functional analyses for *SpFmo3* and *SpSult* in this work failed to show any morphological abnormalities in the knockdown embryos. For *SpSult* it is believed that the experiment was unsuccessful because the annotation provided by the genome was incorrect at that time and has since been updated. For *SpFmo3* there are two possibilities why the experiment did not result in phenotypic knockdown. One possible reason is that the knockdown worked, however the other sea urchin *fmo* members were able to compensate for the loss. Another possibility is that the antisense oligo was not completely effective in blocking translation of the mRNA.

In order to verify that the protein synthesis was indeed blocked, the -100 to +100 *SpFmo3* region will be fused in frame with green fluorescent protein (GFP). The *SpFmo3*::GFP fusion will be injected into eggs either with a) alone or with b) the antisense morpholino or c) with a control morpholino corresponding to a sequence not present in sea urchin. If there is no GFP expression in case b) but there is in case a) and c), then it is known that the antisense morpholino is reaching its target.

Future Directions

The sea urchin GRN for endo-mesoderm specification has over 40 genes as a result of large-scale gene perturbation analyses, computational analyses, genomic data, cis-regulatory analyses and developmental embryology studies (Reviewed in Davidson et al., 2002). The network sub-routine underlying pigment cells development, as the ones required for other SMC types, are currently largely uncharacterized. Other parts of the GRN for endo-mesoderm specification have been extensively characterized but for the most part only in term of number of genes involved and description of their regulatory interactions, that could be direct or indirect. A considerable amount of research remains to identify the direct gene interactions and therefore resolve the structure of this gene network. Direct gene interactions at each node of the network can be experimentally confirmed at the genomic sequence level by cis-regulatory analyses. In the case of pigment cells, future studies of the promoter architecture of *SpPks*, *SpFmo* and *SpSult* will allow the identification of most, if not all, of the direct upstream regulators of this putative differentiation gene battery. The accessibility of the recently published sea urchin genome sequence (Elsik et al., 2006) combined with several computational tools will now be able to accelerate the progression of identifying putative cis-regulatory modules and transcription factor DNA-binding sites.

REFERENCES

Akiyama, Y., Toshihiko, H., Poole, A.M., and Hotta, Y. (1996). The gcm-motif: A novel DNA-binding motif conserved in *Drosophila* and mammals. *Proc. Natl. Acad. Sci. USA* 93: 14912 – 14916.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. <u>Molecular Biology of the Cell</u>. New York and London: Garland Science, 2002.

Amore, G. and Davidson, E.H. (2006). Cis-Regulatory control of cycylophilin, a member of the ETS-DRI skeletogenic gene battery in the sea urchin embryo. *Developmental Biology* 293: 555 – 64.

Anson-Cartwright, L., Dawson, K., Holmyard, D., Fisher, S.J., Lazzarini, R.A., and Cross, J.C. (2000). The glial cells missing-1 protein is essential for branching morphogenesis in the chorioallantoic placenta. Nature Genetics 25: 311 – 314.

Arenas-Mena, C., Cameron, A.R., and Davidson, E.H. (2000). Spatial expression of Hox cluster genes in the ontogeny of a sea urchin. *Development* 127: 4631 – 4643.

Austin, M.B., Saito, T., Bowman, M.E., Haydock, S., Kato, A., Moore, B.S., Kay, R.R., and Noel, J.P. (2006). Biosynthesis of Dictyostelium discoideum differentiation-inducing factor by a hybrid type I fatty acid-type III polyketide synthase. *Nat. Chem. Biol.* 2: 494 – 502.

Bapteste, E., Susko, E., Leigh, J., MacLeod, D., Charlebois, R.L, and Doolittle, W.F. (2005). Do orthologous gene phylogenies really support tree-thinking? *BMC Evol Biol.* 5: 33.

Black, D.L. (2000). Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology. *Cell* 103: 367 – 370.

Blanchard, R.L., Freimuth, R.R., Buck, J., Weinshilboum, R.M., and Coughtrie, M.W. (2004). A proposed nomenclature system for the cytosolic sulfotransferase (SULT) superfamily. *Pharmacogenetics* 14: 199 – 211.

Bohm, B. <u>Introduction of Flavonoids</u>. Singapore: Harwood Academic Publishers, 1998.

Bourlat, S.J., Juliusdottir, T., Lowe, C.J., Freeman, R., Aronowicz, J., Kirschner, M., Lander, E.S., Thorndyke, M., Nakano, H., Kohn, A.B., Heyland, A., Moroz, L.L., Copley, R.R., and Telford, M.J. (2006). Dueterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature* 444: 85 – 88.

Burke, R.D. (1978). The structure of the nervous system of the pluteus larva of Strongylocentrotus purpuratus. *Cell Tissue Res.* 191: 233 – 247.

Burke, R.D. and Alvarez, C.M. (1988). Development of the esophageal muscles in embryos of the sea urchin Strongylocentrotus purpuratus. Cell Tissue Res. 252: 411 – 417.

Calestani, C., Rast, J.P. and Davidson, E.H. (2003). Isolation of pigment cell specific genes in the sea urchin embryo by differential macroarray screening. *Development* 130: 4587 – 4596.

Cameron, R.A., Fraser, S.E., Britten, R.J. and Davidson, E.H. (1991). Macromere cell fates during sea urchin development. *Development* 113: 1085 – 1091.

Cameron, R.A., Mahairas, G., Rast, J.P, Martinez, P., Biondi, T.R., Swartzell, S., Wallace, J.C., Poutska, A.J., Livingston, B.T., Wray, G.A., Ettensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H., and Hood, L. (2000). A sea urchin genome project: Sequence scan, virtual map, and additional resources. *PNAS* 97: 9514 – 9518.

Cashman, J.R. (2000). Human flavin-containing monooxygenase: Substrate specificity and role in drug metabolism. *Curr. Drug Metab.* 1: 181 – 191.

Castle, J., Garrett-Engele, P., Armour, C.D., Duenwald, S.J., Loerch, P.M., Meyer, M.R., Schadt, E.E., Stoughton, R., Parrish, M.L., Shoemaker, D.D., and Johnson, J.M.(2003). Optimization of oligonucleotide arrays and RNA amplification protocols for analysis of transcript structure and alternative splicing. *Genome Biol.* 4: 66.

Castoe, T.A., Stephens, T., Noonan, B.P, and Calestani, C. (2007). A novel group of type I polyketide synthases (PKS) in animals and the complex phylogenomics of PKSs. *Gene*, In Press.

Chou, H.C., Lang, N.P. and Kadlubar, F.F. (1995). Metabolic activation of the *N*-hydroxy derivative of the carcinogen 4-aminobiphenyl by human tissue sulfotransferases. *Carcinogenesis* 16: 413 – 417.

Clark, T.A., Sugnet, C.W., and Ares Jr., M. (2002). Genomewide Analysis of mRNA Processing in Yeast Using Slicing-Specific Microarrays. *Science* 296: 907 – 910.

Davidson, E.H., Cameron, R.A. and Ransick, A. (1998). Specification of cell fate in the sea urchin embryo: summary and some proposed mechanisms. *Development* 125: 3269 – 3290.

Davidson, E.H., McClay, D.R., and Hood, L. (2003). Regulatory gene networks and the properties of the developmental process. *PNAS* 100: 1475 – 1480.

Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otim. O., Brown, C.T., Livi, C.B., Lee, P.Y., Revilla, R., Rust, A.G., Pan, Z., Schilstra, M.J., Clark, P.J.C., Arnone, M.I., Rowen, L., Cameron, A., McClay, D.R., Hood, L., and Bolouri, H. (2002). A Genomic Regulatory Network for Development. *Science* 295: 1669 – 1678.

Dehal, P., Satou, Y., Campbell, R.K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D.M., Harafuji, N., Hastings, K.E.M., Ho, I., Hotta, K., Huang, W., Kawasima, T., Lemaire, P., Martinez, D., Meinertzhagen, I.A., Necula, S., Nonaka, M., Putnam, N., Rash, S., Saiga, H., Satake, M., Terry, A., Yamada, L., Wang, H., Awazu, S., Azumi, K., Boore, J., Branno, M., Chin-box, S., DeSantis, R., Doyle, S., Francino, P., Keys, D.N., Haga, S., Hayashi, H., Hino, K., Imai, K.S., Inaba, K., Kano, S., Kobayashi, K., Kobayashi, M., Lee, B., Makabe, K.W., Manohar, C., Matassi, G., Medina, M., Mochizuki, Y., Mount, S., Morishita, T., Miura, S., Nakayama, A., Nishizaka, S., Nomoto, H., Ohta, F., Oishi, K., Rigoutsos, I., Sano, M., Sasaki, A., Sasakura, Y., Shoguci, E., Shin-I, T., Spagnuolo, A., Stainier, D., Suzuki, M.M., Tassy, O., Takatori, N., Tokuoka, M., Kasumi, Y., Yoshizaki, F., Wada, S., Zhang, C., Hyatt, P.D., Larimer, F., Detter, C., Doggett, P., Glavina, T., Hawkins, T., Richardson, P., Lucas, S., Kohara, Y., Levine, M., Satoh, N., and Rokhsar, D.S. (2002). The Draft Genome of *Ciona intestinalis*: Insights into Chordate and Vertebrate Origins. *SCIENCE* 298: 2157 – 2167.

Delsuc, F., Brinkmann, H., Chourrout, D, Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439: 965 – 968.

Dolphin, C.T., Beckett, D.J., Janmohamed, A., Cullingford, T.E., Smith, R.L., Shephard, E.A. and Philips, I.R. (1998). The flavin-containing mono-oxygenase 2 gene (FMO2) of humans, but not of other primates, encodes a truncated, nonfunctional protein. J. Biol. Chem. 273: 30599 – 30607.

Dolphin, C.T., Cullingford, T.E., Shephard, E.A., Smith, R.L. and Philips, I. R. (1996). Differential developmental and tissue-specific regulation of expression of the genes encoding three members of the flavincontaining monooxygenase family of man, FMO1, FMO3 and FMO4. *Eur. J. Biochem.* 235(3): 683 – 689.

Dunn, R.T., and Klaassen, C. D. (1998). Tissue-specific expression of rat sulfotransferase messenger RNAs. *Drug Metab. Dispos.* 26: 598 – 604.

Elsik, C.G., Worley, K.C., Zhang, L., Milshina, N.V., Jiang, H., Reese, J.T., Childs, K.L., Venkatraman, A., Dickens, C.M., Weinstock, G.M., and Gibbs, R.A. (2006). Community annotation: Procedures, protocols, and supporting tools. *Genome Resources* 16: 1329 – 1333.

Ettensohn, C.A. and Ruffins, S.W. (1993) Mesodermal cell interactions in the sea urchin embryo: Properties of skeletogenic secondary mesenchyme cells. *Development* 117: 1275 – 1285.

Ettensohn, C.A., Wessel, G.M. and Wray, G. editors (2004). Development of Sea Urchins, Ascidians and Other Non-Vertebrate Deuterostomes: An Experimental Analysis. Academic Press

Falany, C.N. (1997). Enzymology of human cytosolic sulfotransferases. *FASEB J.* 11: 206 – 216.

Fox, D.L., and Scheer, B.T. (1941). Comparative studies of the pigments some Pacific Coast echinoderms. *Biol. Bull.* 80: 441 – 455.

Gasser, R. (1996). The flavin-containing monooxygenase system. *Exp. Toxic. Pathol.* 48: 467 – 470.

Gibson, A.W. and Burke, R.D. (1985). The origin of pigment cells in embryos of the sea urchin Strongylocentrotus purpuratus. *Dev. Biol.* 107: 414 – 419.

Gilbert, S.F. (2000). Developmental Biology, 6th Edition. Massachusettes: Sinauer Associaties, Inc.

Gras, H. and Weber, W. (1977). Light-induced alterations in cell shape and pigment displacement in chromatophores of the sea urchin Centrostephanus longispinus. Cell. Tiss. Res. 182: 165 – 176.

Griffiths, M. (1965). A study of the synthesis of naphthaquinone pigments by the larvae of two species of sea urchins and their reciprocal hybrids. *Dev. Biol.* 11: 443 – 447.

Gross, F., Luniak, N., Perlova, O., Gaitatzis, Jenke-Kodama, H., Gerth, K., Gottschalk, D., Dittmann, E., Müller. (2006). Bacterial type III polyketide synthases: phylogenetic analysis and potential for the production of novel secondary metabolites by heterologous expression in pseudomonads. *Archives of Microbiology* 185: 28 – 38.

Gunther, T., Chen, Z.F., Kim, J., Priemel, M., Rueger, J.M., Amling, M., Moseley, J.M., Martin, T.J., Anderson, D.J., and Karsenty, G. Genetic ablation of parathyroid glands reveals another source of parathyroid hormone. *Nature* 406: 199 – 203.

Gustafson, T. and Wolpert, L. (1963). Studies on the cellular basis of the morphogenesis in the sea urchin embryo. *Exp. Cell Res.* 29: 561 – 582.

Harkey, M.A., Whiteley, H.R., and Whiteley, A.H. (1992). Differential expression of the msp130 gene among skeletal lineage cells in the sea urchin embryo: a three dimensional in situ hybridization analysis. *Mech Dev.* 37: 173 – 184.

Her, C., Kaur, G.P., Athwal, R.S. and Weinshilboum, R.M. (1997). Human sulfotransferase SULT1C1: cDNA cloning, tissue-specific expression, and chromosomal localization. *Genomics* 41: 467 – 470.

Hildebrandt, M.A., Salavaggione, O.E., Martin, Y.N., Flynn, H.C., Jalal, S., Wieben, E.D., Weinshilboum, R.M. (2004). Human SULT1A3 pharmacogenetics: gene duplication and functional genomic studies. *Biochem. Biophys. Res. Commun.* 321: 870 – 878.

Hines, R.N., Cashman, J.R., Philpot, R.M., Williams, D.E. and Ziegler, D.M. (1994). The mammalian flavin-containing monooxygenases: molecular characterization and regulation of expression. *Toxicol. Appl. Pharmacol.* 125: 1 – 6.

Hopwood, D.A. (1997). Genetic contributions to understanding polyketide synthases. *Chem. Rev.* 97:2465 – 2497.

Hopwood, D.A. and Sherman, D.H. (1990). Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu. Rev. Genet.* 24: 37 – 66.

Hörstadius, S. (1973). Experimental Embryology of Echinoderm. Oxford: Clarendon Press.

Hosoya, T., Takizawa, K., Nitta, K., Hotta, Y. (1995). *Glial cells missing*: A binary switch between neuronal and glial determination in drosophila. *Cell* 82: 1025 – 1036.

Hu, G.K., Madore, S.J., Moldover, B., Jatkoe, T., Balaban, D., Thomas, J., and Wang, Y. (2001). Predicting Splice Variant from DNA Chip Expression Data. *Genome Research* 11: 1237 – 1245.

Hutchinson, C.R. (1999). Microbial polyketide synthases: more and more prolific. *Proc Natl Acad Sci U S A* 96(7):3336 – 3338

Jain, R., Rivera, M.C., and Lake, J.A. (1999). Horizontal gene transfer among genomes: The complexity hypothesis. *PNAS* 96(7): 3801 – 3806.

Jeffery, W.R., Strickler, A.G., and Yamamoto, Y. (2004). Migratory neural crest-like cells form body pigmentation in a urochordate embryo. *Nature* 431: 696 – 699.

Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R., Shoemaker, D.D. (2003). Genome-Wide Survey of Human Alternative Pre-mRNA Splicing with Exon Junction Microarrays. *Science* 302: 2141 – 2144.

Jones, B.W., Fetter, R.D., Tear, G., and Goodman, C.S. (1995). *glial cells missing*: a genetic switch that controls glial versus neuronal fate. *Cell* 82: 1013 – 1023.

Keating, T.A., and Walsh, C.T. (1999). Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis. *Curr Opin Chem Biol.* 3(5):598 – 606.

Khosla, C., Gokhale, R.S., Jacobsen, J.R., and Cane, D.E. (1999). Tolerance and specificity of polyketide synthases. *Annu Rev Biochem.* 68:219 – 253.

Kindler, S., Wang, H., Richter, D., and Tiedge, H. (2005). RNA TRANSPORT AND LOCAL CONTROL OF TRANSLATION. *Annual Review of Cell and Developmental Biology* 21: 223 – 245.

Kominami, T., Takata, H. and Takaichi, M. (2001). Behavior of pigment cells in gastrula-stage embryos of Hemicentrotus pulcherrimus and Scaphechinus mirabilis. *Dev. Growth Differ.* 43: 699 – 707.

Koukouritaki, S.B., Simpson, P., Yeung, C.K., Rettie, A.E. and Hines, R.N. (2002). Human Hepatic Flavin-containing monooxygenases 1 (FMO1) and 3 (FMO3) developmental expression. *Pediatr. Res.* 51, 236 – 243.

Kuhn, R. and Wallenfells, K. (1940). Echinochrome als prosthetische gruppen hochmolekularern symplexe in den eiern von Arbacia pustulosa. Ber. Dt. Chem. Ges. 73: 458 – 464.

Kwon, H., Smith, W.C., Scharon, A.J., Hwang, S.H., Kurth, M.J., Shen, B. (2002). C-O Bond Formation by Polyketide Synthases. *Science* 297: 1327 – 1330.

Lattard, V., Zhang, J., and Cashman, J.R. (2004). Alternative Processing Events in Human *FMO* Genes. *Mol Pharacol.* 65: 1517 – 1525.

Lederer, E. (1952). Sur les pigments napthoquinoniques des épines et du test des oursins *Paracentrotus lividus* et *Arbacia pustulosa. Biochim. Biophys.* 9: 92 – 101.

Levine, M. and Davdison, E.H. (2005). Gene regulatory networks for development. *PNAS* 102*:* 4936 – 4942.

Logan, C.Y. and McClay, D.R. (1997). The allocation of early blastomeres to the ectoderm and endoderm is variable in the sea urchin embryo. *Development* 11: 2213 – 2223.

Logan, C.Y. and McClay, D.R. 1999. Lineages that give rise to endoderm and mesoderm in the sea urchin embryo. In S. A. Moody (ed.), Cell Lineage and Determination. Academic Press, New York, pp. 41–58.

Lopez, A.J. (1998). Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. *Annu. Rev. Genet.* 32: 279 – 305.

McClay, D.R., Peterson, R.E., Range, R.C., Winter-Vann, A.M. and Ferkowicz, M.J. (2000). A micromere induction signal is activated by beta-catenin and acts through notch to initiate specification of secondary mesenchyme cells in the sea urchin embryo. *Development* 127: 5113 – 5122.

McLendon, J.F. (1912). Echinochrome, a red substance in sea urchins. *J. Biol. Chem.* 11: 435 – 441.

Minchin, R.F., Ilett, K.F., Teitel, C.H., Reeves, P.T., and Kadlubar, F.F. (1992). Direct O-acetylation of N-hydroxy arylamines by acetylsalicylic acid to form carcinogen-DNA adducts. *Carcinogenesis* 13: 663 – 667.

Moore, B.S. and Hopke, J.N. (2001). Discovery of a new bacterial polyketide biosynthetic pathway. *Chembiochem* 2: 35 – 38.

Nemer, M., Rondinelli, E., Infante, D., and Infante, A.A. (1991). Polyubiquitin RNA characteristics and conditional induction in sea urchin embryos. *Developmental Biology* 145: 255 – 265.

Oda, H., Wada, H., Tagawa, K., Akiyama-Oda, Y., Satoh, N., Humphreys, T., Zhang, S., and Tsukita, S. (2002). A novel amphioxus cadherin that localizes to epithelial adherens junctions has an unusual domain organization with implications for chordate phylogeny. *Evolution and Devleopment* 4: 426 – 434.

Ptashne, M. (1986). Gene regulation by proteins acting nearby and at a distance. *Nature* 322: 697 – 701.

Pugh, E.L. and Wakil. S.J. (1965). Studies on the mechanism of fatty acid synthesis. XIV. The prosthetic group of acyl carrier protein and the mode of its attachment to the protein. *J. Biol. Chem.* 240: 4727 – 4733.

Ransick, A. and Davidson, E.H. (2006). Cis-regulatory processing of Notch signaling input to the sea urchin glial cells missing gene during mesoderm specification. *Dev Biol.* 297: 587 – 602.

Ransick, A., Ernst, S., Britten, R.J. and Davidson, E. H. (1993). Whole mount in situ hybridization shows Endo 16 to be a marker for the vegetal plate territory in sea urchin embryos. *Mech. Dev.* 42: 117 – 124.

Ransick A, Rast JP, Minokawa T, Calestani C, Davidson EH. (2002) New early zygotic regulators expressed in endomesoderm of sea urchin embryos discovered by differential array hybridization. *Dev Biol.* **246**(1):132 - 147.

Rast, J.P., Amore, G., Calestani, C., Livi, C.B., Ransick, A., and Davidson, E.H. (2000). Recovery of developmentally define gene sets from high-density cDNA macroarrays. Developmental Biology 228: 270 – 286.

Richardson, A.O. and Palmer, J.D. (2007). Horizontal gene transfer in plants. *Journal of Experimental Botany* 58(1): 1 - 9.

Ruffins, S.W. and Ettensohn, C.A. (1993) A clonal analysis of secondary mesenchyme cell fates in the sea urchin embryo. *Dev. Biol.* 60: 285 – 288.

Ruffins, S.W. and Ettensohn, C.A. (1996) A fate map of the vegetal plate of the sea urchin (Lytechinus variegatus) mesenchyme blastula. *Development* 122: 253 – 263.

Schaefer, B.C. (1995). Revolutions in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. *Anal. Biochem.* 227(2): 255 – 273.

Schliwa, M. and Bereiter-Hahn, J. (1973). Pigment movements in fish melanophores: morphological and physiological studies. III. The effects of colchicine and vinblastine. *Z. Zellforsch.* 147: 127 – 148.

Schreiber, J., Sock, E., Wegner, M. (1997). The regulator of early gliogenesis *glial cells missing* is a transcription factor with a novel type of DNA-binding domain. *PNAS* 94: 4739 – 4744.

Service, M. and Wardlaw, A.C. (1984). Echinochrome-A as a bactericidal substance in the coelomic fluid of Echinus esculentus (L.). *Comp. Biochem. Physiol.* 79B: 161 – 165.

Shen, B. Biosynthesis of Aromatic Polyketides. Berlin: Springer, 2000.

Sherwood, D.R. and McClay, D.R. (1997). Identification and localization of a sea urchin Notch homologue: Insights into vegetal plate regionalization and Notch receptor regulation. *Development* 124: 3363 – 3374.

Sherwood, D.R. and McClay, D.R. (1999). LvNotch signaling mediates secondary mesenchyme specification in the sea urchin embryo. *Development* 126: 1703 – 1713.

Shoemaker, D.D., Schadt, E.E., Armour, C.D., He, Y.D., Garrett-Engele, P., McDonagh, P.D., Loerch, P.M., Leonardson, A., Lum, P.Y., Cavet, G., Wu, L.F., Altschuler, S.J., Edwards, S., King, J., Tsang, J.S., Schimmack, G., Schelter, J.M., Koch, J., Ziman, M., Marton, M.J., Li, B., Cundiff, P., Ward, T., Castle, J., Krolewski, M., Meyer, M.R., Mao, M., Burchard, J., Kidd, M.J., Dai, H., Phillips, W., Linsley, P.S., Stoughton, R., Scherer, S., and Boguski, M.S. (2001). Experimental annotation of the human genome using microarray technology. *Nature* 409: 922 – 927.

Soltis, P.S., and Soltis, D.E. (2003). Applying the bootstrap in phylogeny reconstruction. *Stats. Sci.* 18: 256 – 267.

Staunton, J. and Weissman, K.J. (2001). Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* 18: 380 – 416.

Summers, R.G. Morrill, J.B, Leith, A., Marko, M., Piston, D.W., and Stonebraker, A.T. (1993). A stereometric analysis of karyogenesis, cytokinesis, and cell arrangements during and following fourth cleavage period in the sea urchin, *Lytechinus variegatus Dev. Growth Diff.* 35: 41-58.

Tamboline, C.R. and Burke, R.D. (1992). Secondary mesenchyme of the sea urchin embryo: Ontogeny of blastocoelar cells. *J. Exp. Zool.* 262: 51 – 60.

Thornton, J.W. and DeSalle, R. (2000). GENE FAMILY EVOLUTION AND HOMOLOGY: Genomics Meets Phylogenetics. *Annual Review of Genomics and Human Genetics* 1: 41 – 73.

Tyler, A. (1939). Crystalline echinochrome and spinochrome: their failure to stimulate the respiration of eggs and of sperm of *Strongylocentrotus*. *Proc. Natl. Acad. Sci. U.S.* 24: 523 – 528.

Yoon, K and Gaiano, N. (2005). Notch signaling in the mammalian central nervous system: insights from mouse mutants. *Neuroscience* 8: 709 – 715.

Yuh, C.H., Brown, C.T., Livi, C.B., Rowen, L., Clarke, P.J.C., and Davidson, E.H. (2002). *Developmental Biology* 246: 148 – 161.

Varin, L., and lbrahim, R.K. (1989) Partial purification and characteri-zation of three flavonol-specific sulfotransferases from *Flaveria chlorae-folia. Plant Physiol.* 90: 977 – 981.

Wada, H., Okuyama, M., Satoh, N., and Zhang, S. (2006). Molecular evolution of fibrillar collagen in chordates, with implications for the evolution of vertebrate skeletons and chordate phylogeny. *Evolution and Development* 8: 370 – 377.

Wang, H., Hubbell, E., Hu, J., Mei, G., Cline, M., Lu, G., Clark, T., Siani-Rose, M.A., Ares, M., Kulp, D.C., and Haussler, D. (2003). Gene structure-based splice variant deconvolution using a microarray platform. Bioinformatics 19: 315 – 322.

Weber, W. and Dambach, M. (1974). Light-sensitivity of isolated pigment cells of the sea urchin Centrostephanus longispinus. *Cell Tiss. Res.* 148: 437 – 440.

Weber, W. and Gras, H. (1980). Ultrastructural observations on changes in cell shape in chromatophores of the sea urchin Centrostephanus longispinus. *Cell Tiss. Res.* 206: 21 – 33.

Weinshilboum, R.M., Otterness, D.M., Aksoy, I.A., Wood, T.C., Her, C. and Raftogianis, R.B. (1997). Sulfation and sulfotransferases 1: Sulfotransferase molecular biology: cDNAs and genes. *FASEB J.* 11: 3 – 14.

Wise, G.E. (1969). Ultrastructure of amphibian melanophores after light-dark adaptation and hormonal treatment. J. *Ultrastruct. Res.* 27: 472 – 485.

Wray, G.A. (1999). Introduction to sea urchins. In S. A. Moody (ed.), Cell Lineage and Determination. Academic Press, New York. pp. 3 –9.

Yeakley, J.M., Fan, J., Doucet, D., Luo, L., Wickham, E., Ye, Z., Chee, M.S., and Fu, X. (2002). Profiling alternative splicing on fiber-optic arrays. *Nature Biotechnology* 20: 353 – 358.

Zmasek, C.M. and Eddy, S.R. (2001). ATV: display and manipulation of annotated phylogenetic trees. *Bioinformatics* 17: 383 – 384.