REGENERATION IN THE CTENOPHORE *MNEMIOPSIS LEIDYI*: A TRANSCRIPTOMIC ANALYSIS

By

RACHEL SUSAN SANFORD

A THESIS PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2015

To my parents, Bradley and my cat Tiger

ACKNOWLEDGEMENTS

I thank my parents for the opportunity to be here and the constant support. I would like to thank my fellow laboratory colleagues for our hours at the bench. I would like to thank my supervisors and mentors for the direction they have given me. I thank my friends for all the help they given me to complete this thesis.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

All photos and illustrations were produced by the author.

LIST OF OBJECTS

LIST OF ABBREVIATIONS

DIC	Video differential contrast
EdU	5-ethynyl-2’-deoxyuridine
FDR	False Discovery Rate
FSW	Filtered sea water
Gbps	Giga base pairs, 1,000,000,000 base pairs
HPC	HiPerGator High Performance Computing System
ICBR	Interdisciplinary Center for Biotechnology Research
ICG	Interplate ciliated groove
MlSP	Mnemiopsis leidyi Secretory Peptide, ctenophore specific secretory peptide first discovered in Mnemiopsis
PbSP	Pleurobrachia bachei Secretory Peptide, ctenophore specific secretory peptide first discovered in Pleurobrachia.
PCA	Principle Component Analysis
PGM	Personal Genome Machine, IonTorrent sequencing machine
RNA-seq	RNA sequencing
RPKM	Reads Per Kilobase Per Million mapped reads
SP	Secretory Peptide

Abstract of Thesis Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Master of Science

REGENERATION IN THE CTENOPHORE *MNEMIOPSIS LEIDYI*:
A TRANSCRIPTOMIC ANALYSIS

By

Rachel Susan Sanford
December 2015
Chair: Leonid Moroz
Major: Medical Sciences

Ctenophores' amazing capacity of regeneration has fascinated biologists for centuries. The morphological features of ctenophore regeneration have been documented, but the molecular and cellular components behind this phenomenon have remained a mystery. Here, next generation sequencing technologies and transcriptomic analysis are used to investigate the regeneration dynamics in the ctenophore *Mnemiopsis leidyi*. The resulting data identify multiple signaling pathways that might be involved in ctenophore regeneration. These include evolutionarily conserved pathways, such as Ca^{2+}-dependent and MAP-kinase signaling pathways, that are up regulated during regeneration, as well as genes involved in energetics and cytoskeleton function. The data also show evidence for involvement of dozens of ctenophore specific secretory molecules, their receptors and processing components that are important signal messengers in regeneration. A unique subset of transcription factors were also found to be involved in regeneration which may be upstream regulators of those signaling pathways. In summary, our data indicate that the strategies which ctenophores employ to regenerate use a unique combination of evolutionarily conserved and ctenophore specific signaling components. These data provide novel insights into the mechanisms of regeneration in the earliest branching taxa in Metazoa.

CHAPTER 1
INTRODUCTION

Regeneration in Metazoa

Regeneration, the regrowth or repair of cells, tissues or organs, is widely represented among animal phyla (Fig. 1-1) (1, 2). Regeneration can occur at various levels of biological organization such as whole body, structure (e.g. limb, fin, tail), organ systems (e.g. heart, liver), tissue (e.g. epidermis, gut lining), or cellular (e.g. axon, muscle fiber) (3). Though the capacity for regeneration is widely spread through the animal taxa, it is not consistently represented. Some clades include species that can regenerate their whole body from only a small piece such as the flatworms *Dugesia japonica*, others can only regenerate specific structures like the fiddler crabs limb regeneration. In this body of work I will be focus on tissue regeneration.

There are two general cellular mechanisms through which regeneration can take place, first described by T H Morgan in his account of regeneration in 1901. Epimorphisis involves cell proliferation and formation of a blastema, a mass of undifferentiated cells capable of growth or regeneration, where the structure which is being replaced originates, whereas Morphallaxis involves little to no cell proliferation, but the movement of existing cells to form new tissue (4).

In addition to the phenomenon itself, these variables in regeneration have raised many questions that remain largely unanswered: Which molecules and genes are key mediators of regeneration? Are multiple signaling pathways present in regeneration? Though there are many unanswered questions, the molecular components of regeneration have been studied in many different capacities. The signaling components behind zebrafish fin regeneration (5), the molecular components to limb regeneration

(6), regeneration in vertebrate systems (7), wound repair (8), and even the molecular mechanisms of neuronal regeneration have been investigated (9, 10).

In this body of work, I investigated the genetic regulation of regeneration, utilizing an organism from a basally branched taxa, *Mnemopsis leidyi.* As the molecular components of regeneration have not been previously investigated in this species, I will take both a top-down approach by comparing previously studied regeneration associated genes to data gathered from these series of experiments, as well as a bottom-up approach by exploring which genes are most differentially expressed in the regenerative process.

Ctenophora

The regenerative model organism that will be focused on here is the ctenophore *Mnemopsis leidyi,* (Fig1-2A) an organism from the phylum Ctenophora with stunning regenerative capacities. Ctenophores, also called comb jellies, are gelatinous marine organisms with a global distribution. Recent evidence has shown this phylum is the earliest branch of the Metazoa (11-13), placing them sister to Porifera, Placazoa, Cnidaria and Bilateria. Though there have been many observational studies of *Mnemiopsis's* ability to regenerate very little is known about the molecular and genomic components of this regeneration. The purpose of this body of work is to study these molecular and genomic mechanisms involved in ctenophore regeneration. Studying these mechanisms in ctenophores may serve to provide insights into the shared processes mediating regeneration in diverse model organisms, as their basal position in Metazoa phylogeny suggests that regenerative capacities have early, and potentially conserved evolutionary roots.

Anatomy

Ctenophores vary in size from a few millimeters to over a meter (14). Their name derives from the eight rows of ciliary comb plates, or ctenes, which distinguish them from all other animals (Fig 1-2B). These are the largest known animals to swim with ciliary motion and the comb plates of ctenophores are the largest ciliary structures that have been studied (15). The refraction of light from the comb plates gives ctenophores their characteristic iridescence. The body plan of ctenophores has a major axis in the oral-aboral direction (16), with an aboral sensory organ at one end of this axis and a mouth at the other. These animals are biradially symmetrical with four identical quadrants. Two planes of symmetry run along the oral-aboral pole: one plane of symmetry runs through the tentacles (tentacular axis) and another runs through the pharynx (sagittal axis) (17) (Fig 1-2C). The eight ctene rows are grouped by these subdivisions. The rows that are closest to the tentacles are called the subtentacular comb rows and the rows closest to the pharynx are called the subsagittal comb rows (15). From the aboral sensory organ narrow tracks containing small cilia, called ciliated grooves, run to the first comb plate. In lobate ctenophores, such as *Mnemiopsis*, these ciliated grooves continue between successive comb plates through the whole comb row. Locomotion in ctenophores is accomplished by the synchronous beating of the individual comb plates on each of the comb rows. In lobate ctenophores motion is accomplished through waves of ciliary beating which start at the pacemaker cilia and travel orally (18).

Ctenophores have two tentacles that are covered in specialized cells called colloblasts, which are unique adhesive cells possibly used for trapping prey (19). All known ctenophores are predators and most eat plankton in the water column (20). Their

digestive system consists of a mouth at the oral end of the animal that leads to a flattened and elongated pharynx, which then leads to a stomach. From the stomach there are canals that branch off to the rest of the body. The canals which branch off to the aboral poles lead to anal pores which open to the exterior. Therefore paths for ingestion and egestion are separate and these processes can occur at the same time (21). Most ctenophores have the capacity for bioluminescence, and these light producing cells are generally centered around the ctene rows (22).

The neural system of ctenophores is made up of multiple cell types: nerve cells forming a subepithelial neural net, which make up a polygonal lattice covering the body surface ctenophore (23), mesogleal neural net cells, neural elements in the tentacles and gastrodermal system, and the neural cells in the aboral organ (24). The aboral organ contains a statocyst with a single large statolith, supported by balancing cells, which control the beat frequency of the comb rows. The statolith is comprised of many small cells called lithocytes. A cilia dome encloses the statolith and balancing cells. This organ is connected with the rows of cilia by the four ciliated grooves. Two polar fields run along the sagittal plane (25, 26) (Fig 1-2D).

Regeneration in Ctenophores

Though the biology of ctenophores had been studied by Chun and others in the 1880s (27) their ability to regenerate was not discovered until Mortensen further experimented with these animals in 1915 (28). Initially Motensen suspected the ctenophore genus named at that time "*Lesuereia*" was actually a highly injured ctenophore *Bolinopsis infundibulum*, as the general way of retrieving these animals was to capture them with plankton nets from ships, sometimes in rough water. Motensen caught these animals, some which had lost their lobes completely, and placed them in a

large jar for observation. Over the next week he observed them regenerate into *Bolinopsis*. Subsequently, Motensen performed experiments where he isolated the aboral organ from one animal, which quickly regenerated. He then divided two animals longitudinally and one transversally; three of the four longitudinal halves regenerated into new animals, and the aboral half of the longitudinal injury regenerated successfully. These experiments marked the first evidence for regeneration in ctenophores (28).

Coonfield systematically studied regeneration in *Mnemiopsis leidyi* in 1936. Using a larger sample size he concluded halves, thirds and fourths of *Mnemiopsis* have the capacity to regenerate into new animals while eighths fail to regenerate. The sequence of regeneration followed the order of apical organ then comb rows. He also concluded that the apical organ was a regulating center for regeneration (29). In later experiments, Coonfield posed the question: is the rapid regeneration in *Mnemiopsis* due to rapid movement of cells from the remaining parts of the organ or from migration of non-specified cells? This was tested by removing either small sections of a comb row, an entire comb row, removing both ends of an animal, removing either part of a comb row or an entire comb row. He concluded that the method of regeneration depended on how the injury happens: when a whole comb row or part of a comb row is removed from an animal, the regeneration is due to cells from the remaining parts of the canal or neighboring canals; when the entire canal above where the plates grow is removed the regeneration is due to migrating non-specified cells from the mesoglia (30). Coonfield also conducted an experiment to test for a physiological gradient in regeneration, where he found injury placement along the axis of the animal did not affect the regeneration time and demonstrated that there is no physiological gradient present in *Mnemiopsis*

(31). Studying symmetry in ctenophores, Coonfield observed that reorganization during regeneration is a central aspect of regulation and body symmetry in ctenophores (32).

In 1967 Freeman conducted a study on regeneration in the benthic ctenophore, *Vallicula multiformis*, describing the process of asexual reproduction and regeneration. *Vallicula* can reproduce asexually by segregating pieces of tissue from its periphery, and those pieces then go on to form new animals. This process was observed in a group of 10 animals, placed in separate finger bowls, which were segregated into 183 pieces in a 20-day period, 169 of which differentiated into whole animals. Freeman also experimentally showed the regeneration capacity of this animal by dissecting specimens into as many as 11 different pieces, all of them regenerating a new animal. He states:

"Any part of the animal containing an organ or organs can regenerate any missing organs to form a complete individual" (33) p 75.

Another interesting observation in this study is the presence of what Freeman described as "half-animals." These are defined as animals that have only regenerated one set of tentacles and one apical organ. In asexual reproduction, this occurred 12 times, which represented 1% of the population. Freeman experimentally produced these "half-animals" 30% of the time by cutting *Vallicula* in half through the aboral organ along the sagittal plane. Here Freeman explains that the "half-animal" is an occasion where regeneration has not fully been accomplished, but maintains a semi-stable state where the animal does not complete regenerating (33).

In 1986 a study by Martindale posed the question, when does the ability to recognize and replace missing parts begin in the life history of the ctenophore? Ctenophore embryos lack the ability to regenerate after experimental injury, while adults

regenerate easily. Martindale studied *Mnemiopsis leidyi* to find when their regenerative capabilities arose during development. Previous findings by Chun (27), Fischel (34), and Farfagllo (35) have shown embryos separated while a blastomere at the two and four cell stages results in “half-animals” or “quarter animals.” These “half-animals” from separating the blastomeres on the sagittal axis have four ctene rows (two sets), half an apical organ, and one tentacle. The “quarter-animals” have only two ctene rows. Sometimes, these deficient embryos can replace missing structures in the adult phase. This phenomenon is called “post-regeneration.” Martindale bisected embryos at various times during development to determine the time in development where adult regenerative potential begins. The results indicated that the time of ctene row coordination in embryo development represents the transition point in the ability of *Mnemiopsis* to regenerate normal symmetrical properties (17).

Martindale also found that there were two modes of regeneration in *Mnemiopsis*. The first mode was to regenerate a full component of the structures which were lost at the wound site. The second mode was a two step process which involved the animal healing and behaving as a “half-animal” – that is having half the required organs – for a short period of time then a day or two later the remaining structures regenerating. The first mode was most common (17). In 1996 Martindale and Henry investigated whether the same cell lineage that produces ctene plates during embryogenesis also makes them during “post-generation”, when defective embryos replace missing structures in the adult phase. They destroyed two adjacent e_1 micromeres in a 16-cell-stage embryo of *Mnemiopsis leidyi*, which previous experiments have shown to produce a comb row (25), then labeled the other blastomeres with lineage tracer DiI. The results of this study

show that the m_1 lineage, but not that of the other blastomeres, could produce pluripotent stem cells that are capable of giving rise to novel cell types such as comb rows which they would not normally give rise to in embryogenesis (36).

Henry and Martindale studied the phenomenon of “post-regeneration” in 2000. As previously stated, embryos bisected early in development form stable “half-animals”. Additionally, if one of the four e_1 micromeres is removed the larvae does not form any ctene rows (17), but in this case the missing ctene rows will ultimately be repaired during the larval and adult phases in what is called “post-regeneration”. But not all partial embryos will undergo “post-regeneration.” What are the conditions that an embryo will undergo process? The results of this study show that successful post regeneration can be predicted based on a modified polar coordinate model and the rules of intercalary regeneration (37). The stability of the “half-animals” comes from the idea that positional values are located across from each other in the embryonic and adult body plans are equivalent (38).

Andrilenas and Moroz studied injury to *Bolinopsis infundibulum* and *Pleurobrachia bachei* in 2010. In this study, immunohistochemistry was used to label the signaling molecule FMRF-amide, which appeared in post-injury neurogenesis suggesting its involvement in regeneration. This study also found *Bolinopsis* to be quicker at regeneration than *Pleurobrachia* (39).

In 2011 Alie et al. investigated the presence of *Piwi* and *Vasa* in the somatic stem cells of adult ctenophores. While this experiment was not directly an injury experiment, stem cells are involved in regeneration and provide evidence as to which genes are involved. Evidence that the transposon silencer *Piwi* and the translational

activator *Vasa* are involved in germline determination and maintenance has been shown in classic models (40, 41). Alie studied the expression of *Piwi*, *Vasa* as well as *Bruno* and *PL10* in the ctenophore *Pleurobrachia pileus,* and found expression of all genes in the male and female gametes with the exception of one *Piwi* paralog. In addition, these genes showed similar expression in somatic territories such as tentacle root, comb rows and aboral organ. These matched to EdU (5-ethynyl-2'-deoxyuridine) DNA labeling of stem cell concentration areas in these territories (42).

In 2012, Tamm and colleagues used video differential interference contrast (DIC) microscopy to study the morphology of the regeneration of the comb plates in *Mnemiopsis leidyi*. Directly after dissection of the comb row the wound closed and the distance between the comb rows near the cut increased. He found widening of interplate ciliated groove (ICG), growth of opposing comb rows on opposite site of the ICG, elongation of comb rows, and widening until they merged into a single plate. He also found that the timing of this process was variable between comb plates (43).

Transcriptomics

In this study we used RNA-seq to study the molecular components behind ctenophore regeneration. In RNA sequencing, the current set of transcripts in a cell or tissue, and their expression for that physiological condition is quantified. The transcriptome is the complete set of transcripts in the cell for a particular condition (44). Using transcriptomics this study is taking a broader approach than previously used methods for studying regeneration. Transcriptomics allows for an unbiased sample all of the RNA being produced in the cell or tissue. In previous gene specific approaches, an important gene might be missed simply because there was no prior indication to look for it.

In this study both Ion Torrent and Illumina sequencing were used. The organization of the data in this study will correspond to the sequencing technology.

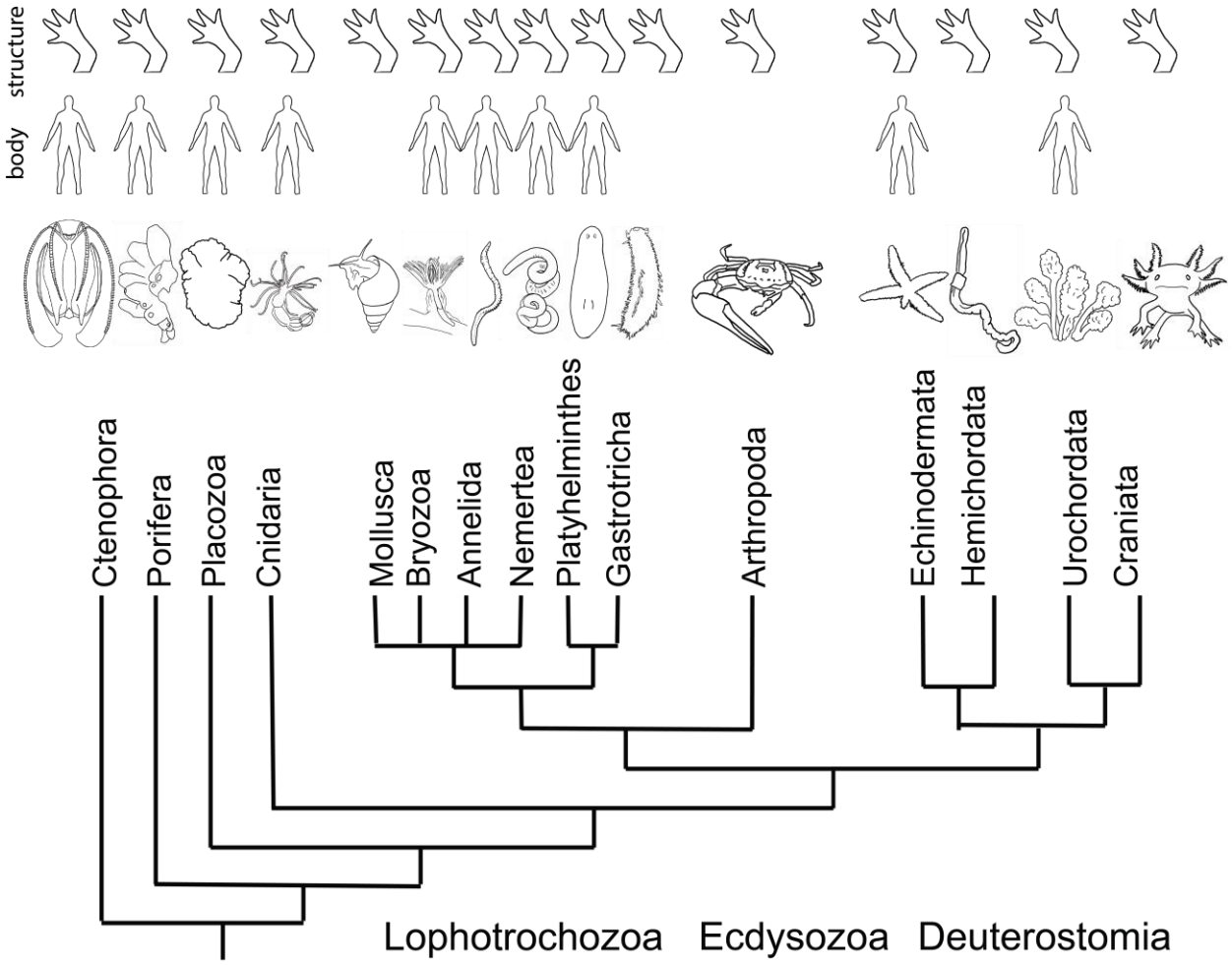


Figure 1-1. Phylogenic representation of regeneration throughout the metazoa. These are all phyla that contain species which has regenerative capabilities. An example of such a species is represented above each phyla. These species are as follows. Basal Metazoans: Ctenophora- *Mnemiopsis leidyi*, Porifera- *Haliclona curacaoensis*, Placozoa- *Trichoplax adhaerens*, Cnidaria- *Nematostella vectensis*. Lophotrochozoa: Mollusca- *Hydrobia ulvae*, Bryozoa- *Plumatella fungosa,* Annelida- *Capitella teleta*, Nemertea- *Lineus pictifrons*, Platyhelminthes- *Dugesia japonica*, Gastrotricha- *Thaumastoderma ramuliferum*. Ecdysozoa: Anthropoda- *Uca pugilator*. Deuterostomia: Echinodermata- *Asterias rubens*, Hemichordata- *Saccoglossus pusilis*, Urochordata- *Pyura spinifera*, Craniata- *Ambystoma mexicanum*. Phylogeny based off of Bely 2010 (2) and modified with data from Moroz et al. 2014 (11), Whelan et al. 2015 (12) and Ryan et al. 2013 (13).

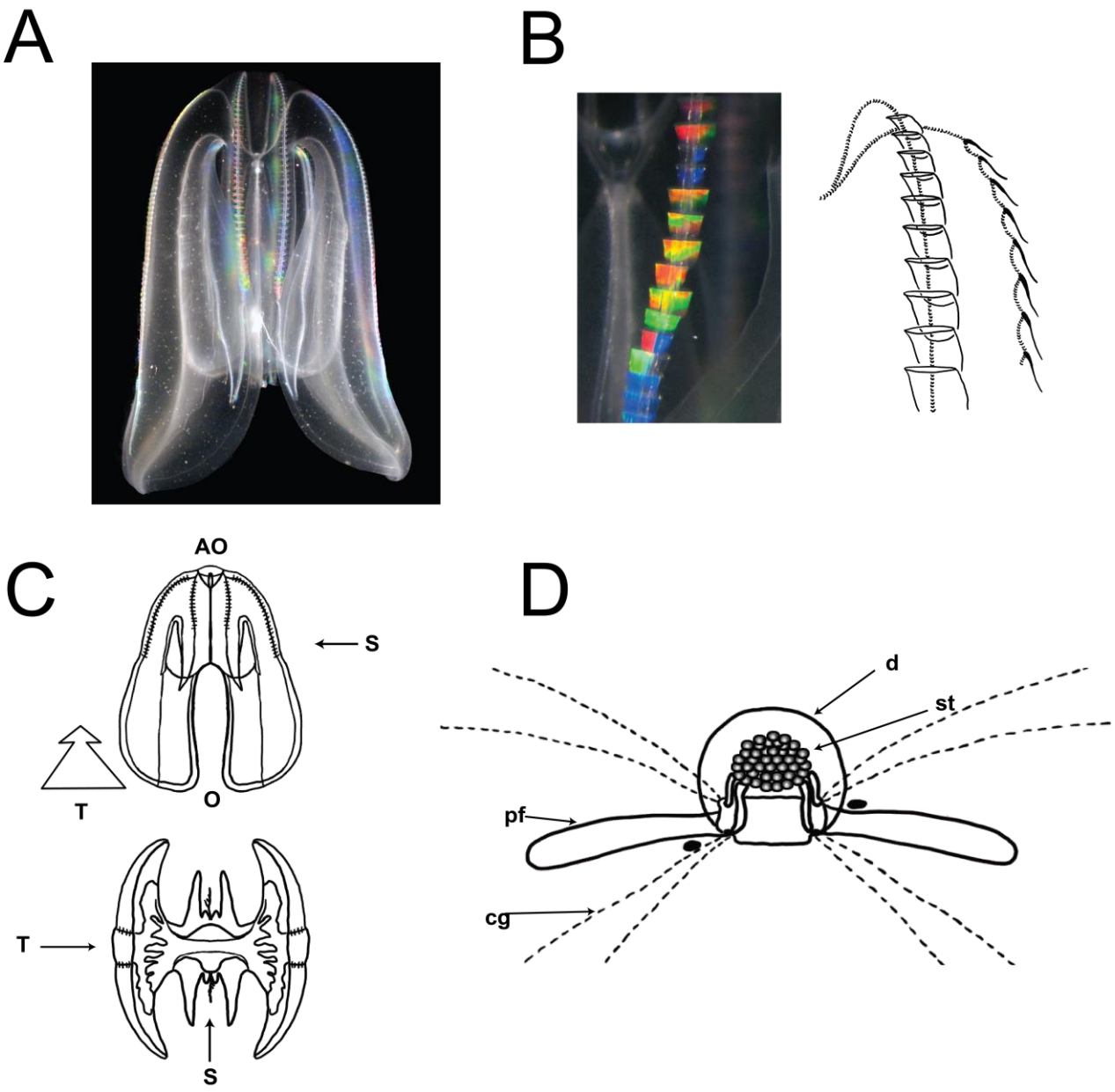


Figure 1-2. Anatomy and introduction to ctenophores. A) Photograph of *Mnemiopsis leidyi* by the author showing the sagittal plane. B) Photograph and diagram by the author of the ctenes or combs which ctenophores derive their name. This ciliary structure is used for locomotion by a coordinated beating of these paddle-like combs along the comb row. Diagram modified from Tamm 2014 (15). C) Biradial body symmetry in ctenophores. S = Sagittal axis; T= Tentacular axis. AO =Aboral pole; O= Oral pole. Modified from Martindale 1986 (17). D) Aboral organ. pf = polar field; cg = ciliated groove; d = dome cilia; st= statolith. Modified from Parker and Haswell 1867 (45).

CHAPTER 2
METHODS

Animal collection and husbandry

Mnemiopsis leidyi were collected from three locations. First, the animals were collected locally in the marina adjacent to the Whitney Laboratory for Biomarine Science (Whitney Lab) in St Augustine, FL located at 9505 Ocean Shore Blvd St. Augustine, FL 32136, *Mnemiopsis leidyi*. These animals were collected using a plastic beaker attached to a long pole, which was used to scoop them out of the water. The second location from which animals were obtained was Woods Hole Oceanographic Institution located at 86 Water St, Woods Hole, MA 02543, also *Mnemiopsis leidyi*. These animals were then shipped to the Whitney Lab.

The ctenophores were housed a few different ways. One of these ways was to place them in large sea tanks at the Whitney Lab with continuous flow of sea water. Under these conditions, they were kept in small floating beakers with mesh sides to reduce damage to the ctenophores; however, animals could only be kept this way for a few days. Another way ctenophores were housed was in a separate small tank designed to care for jellyfish. The tank was obtained from the company JellyfishArt (http://www.jellyfishart.com; Desktop Jellyfish Tank). The tank is a flat cylinder of Plexiglas with an opening on the top of the curve. It rests on a stand and has a slow bubbler which makes the water flow slowly on the bottom of the tank. This bubbler sits under artificial rocks under a plexiglass covering. This tank has a very slow flow of water around the tank which reduces damage to the enclosed animals. Sea water from the Whitney Lab was UV treated before being placed into the tank in order to reduce bacteria. Water was changed every two days to maintain ctenophore health. On the

days the water was not changed the top layer of water was removed to reduce the mucus build up. Lastly, we utilized a large slow flowing tank designed for the housing of ctenophores, jellyfish or other fragile marine creatures made by partners of the Whitney Lab. This large Plexiglas tank allowed constant sea water flow but it is designed where there is very slow suction at the bottom of the tank and slow input at the bottom as well. The tall dimension allows vertical movement of ctenophores which our observations suggest is beneficial to their maintenance. Ctenophores were typically housed for two or three days after they were obtained, prior to being used for experiments.

Planktonic food for the animals was collected by plankton net. A large plankton sample was collected in a five gallon bucket. To condense this sample and reduce the amount of debris introduced into the ctenophore habitat, we placed a bright light on the side of the bucket, waited approximately 30 minutes for the phototaxic plankton to concentrate and drew these up with a turkey baster. This water was then filtered through an 80 mm mesh, which will pass water but retain large plankton. Another food used was Jellyfish Food, obtained from JellyfishArt (http://www.jellyfishart.com; 30mL Jellyfish Food); this product is made up of several plankton ingredients, up to 800 microns. Animals were fed one to two times a day.

Dissection procedures

Prior to dissection of tissue, animals were rinsed three times with filtered sea water (0.2 micron filter unit Fisher cat no. 09-740-26A) in glass dishes. Animals were then transferred to large petri dishes (Falcon Bacteriological Petri Dishes with Lid; Fisher cat no. 08-757-100D) for dissection and viewed under a dissecting scope. Dissections were performed using high quality dissection instruments.

For each injury experiment, an injury was performed, followed by a specified delay before the regenerating tissue around the injury site was dissected. The dissected tissue was then placed into RNA lysis buffer from the Ambion RNAaqueous kit (Ambion™ RNAqueous-Micro Total RNA Isolation Kit; cat no. AM1931 or Ambion™ RNAqueous Total RNA Isolation Kit; cat no. AM1912), dissolved and frozen at -20°C until RNA could be isolated. Control tissue was extracted directly from a healthy, previously untouched animal. The injury site varied for each experiment, as well as the number of animals used. Each experiment is described below, as it relates to the sequencing platform used. See Figure 2-1 for a visual representation of the dissection procedures.

Whole Body Injury Juvenile Ion Torrent

We performed RNA-Seq capture with whole body injury. Two animals were used, each injured each in multiple places. For one animal, tissue was dissected into RNA lysis buffer at 1 hour post-injury, dissolved and frozen at -20º C. For the second animal tissue was placed into RNA lysis buffer at 2 hours post-injury, dissolved and frozen at -20º C.

Aboral Organ Removal Ion Torrent

In this experiment four animals were used. The aboral organ was dissected out of each animal and discarded. One hour after dissection the surrounding regenerating tissue from the first animal was placed into RNA lysis buffer, dissolved and frozen at -20º. This was repeated at 2 hours post-injury for a second animal. The other two animals were returned to flowing sea water. At 24 hours one of the remaining animals was removed and the tissue dissected into RNA lysis buffer, and this repeated at 48 hours after initial dissection.

Epithelial Injury Ion Torrent

Three animals were used in this experiment. Injury was performed on only epithelial tissue. The first sample of regenerating epithelial tissue was collected 10 minutes after initial injury and placed into lysis buffer. The other animals both had control samples taken from them. Both of them had epithelial tissue injured. For one animal, regenerating tissue was dissected 1 hour post-injury and placed into lysis buffer. For the other animal, regenerating tissue was dissected 2 hours post-injury and placed into lysis buffer.

Epithelial Injury Illumina

One animal was used in this experiment. Control tissue was first removed and placed into lysis buffer. Next, the animal was injured only in epithelial tissue, which was placed into a dish of filtered sea water. At intervals of 30 min, 1 hour, and 2 hours samples were taken in triplicate from this animal.

Transcriptome Sequencing

All samples were processed for transcriptome sequencing. Figure 2-2 shows the transcriptome sequencing process.

RNA Isolation

The samples were frozen in Ambion RNAaqueous lysis buffer, either micro for small amounts of tissue or normal for larger amounts of tissue. These lysis buffers are compatible with QIAGEN RNA isolation kits and they have the advantage that they can be frozen and stored. The kits used here to isolate the RNA are QIAGEN RNeasy Micro Kit (cat no. 74004) and RNeasy Mini Kit (cat no. 74104). The samples were thawed to room temperature and centrifuged. The supernatant was transferred to a microcentrifuge tube and 1 volume 70% EtOH was added. This mixture of chaotropic

salts and ethanol was transferred to the spin column, which bound the RNA to the silica membrane. This was centrifuged and a series of washes removed any impurities from the membrane, leaving only RNA. Finally we used Nuclease-Free water (cat no. AM9937) to elute the sample that unbound the RNA from the column. The quality and concentration of this eluted RNA was checked on the Agilent 2200 TapeStation, and then frozen at -80º C until use.

Library Construction

All pipetting for library construction was done with Rainin Classic Pipettes PR-20, PR-100, PR-200, and PR-1000 and Rainin low retention tips RT-L20, RT-L200, and RT-L1000. All tubes used in library construction were LoBind, either 1.5 mL Eppendorf LoBind Microcentrifuge Tubes 022431021 or 0.5 mL Eppendorf LoBind Microcentrifuge Tubes 022431005.

cDNA synthesis

To make the cDNA we used the SMARTscribe Reverse Transcriptase (Fisher cat no. NC9877888) and Ultrapure PCR dNTP mix (Fisher cat no NC9287432). First a master mix was made with 5X first strand buffer, 20 mM DDT, 10 mM dNTPs, CapOligo and RNase Inhibitor. In a separate PCR tube 2 ng – 1ug total of RNA was added to 1 ul of CapT30 10uM primer. This was incubated at 72º C for 3 minutes. The SMARTscribe reverse transcriptase was added to the master mix and after the 3 minutes, the master mix was added to the PCR tube. The tube incubated at 42º C for an hour and a half to prepare the cDNA. Next, to make PCR copies of the cDNA we used Advantage 2 PCR enzyme kit (Fisher cat no NC9581482). We made a master mix using Nuclease-Free water, 10X advantage 2 PCR buffer, 10 mM dNTP mix, 10uM CapPCR, and 50X Advantage 2 polyermase mix and added this to the PCR tube after the initial incubation.

We ran it on a PCR program of 95º C for 1 min, 95º C for 15 sec, 65º C for 30 sec, 68º C for 30 sec, repeated step 2 to 3 18-25 X depending on the initial RNA concentration, then held at 4º C. We checked this on an E-Gel® Precast Agarose Electrophoresis System.

Purification of cDNA

We purified the cDNA using Beckman Coulter Agencourt AMPure XP magnetic beads (Cat no A63881). This involved mixing the cDNA with 1.8X of the beads and incubating the mixture, then letting it incubate on a magnetic rack and removing the supernatant. After washing the beads twice with EtOH the beads are dried and then they can be eluted with Nuclease-Free water.

Fragmentation of DNA

We perform the fragmentation with shearing by using the Covaris M220 Focused-ultrasonicator™. The purified DNA was placed in a 50 mL microTUBE AFA fiber Screw-Cap (Cat no. NC0380760) and sonicated to either 200 bp for Ion Torrent or 400 bp for Illumina.

End repair

The rest of the library preparation procedure was performed with NEBNext Ultra DNA Library Prep Kit for Illumina (Cat no 50591145) or NEBNext Fast DNA Library Prep Set for Ion Torrent (Cat no E6270S). To repair the ends of the fragmented DNA it was mixed with end repair reaction buffer and end repair enzyme mix, then incubated in a thermocycler at 20º C for 30 min and 65º C for 30 min.

Adaptor ligation

To ligate on the specific adaptors needed we performed the steps outlined in the protocols. For the illumina protocol it specified that if the DNA input is less than 100 ng,

to dilute the adaptors provided in the kit 10 fold. This was the case with all of the DNA and we diluted these adaptors with Nuclease-Free water. In the illumina kit, the TA master mix, diluted adaptors, ligation enhancer and the DNA were mixed and this was incubated at 20º C for 15 minutes. Then the user enzyme was added and this was incubated another 15 minutes at 37º C. For the Ion Torrent protocol the T4 Ligase Buffer, T4 Ligase, DNA polymerase, adaptors, and the DNA was mixed and this was incubated at 25º C for 15 minutes followed by 65º C for 5 minutes.

Size selection

The size selection is performed with Beckman Coulter SPRIselect beads (Cat no B23317). We perform a double size selection with a right side ratio of 0.64x and a left side ratio of 0.75x. When the beads were mixed with the sample, the lower the ratio, the higher the size cutoff. As we added fewer beads, larger fragments would bind. For the double size selection we added 0.64x for the right side. This bound fragments above ~400 bps onto the beads and below ~400 were kept in the supernatant, so we kept the supernatant and discarded the beads. Now for the left side we added 0.75x that binds fragments above 200 bps on the beads and below 200 is in the supernatant and we kept the beads. The library we had is now 200 to 400 bps in size.

PCR amplification of final library

The size selected DNA, primers and master mix were added together for the PCR. This was the step to add the NEBNext Multiplex Oligo Index Primers (#E7335 or #E7500) for each library. Care was taken to not contaminate or mix any of the individual barcodes. The initial denaturation was 98 for 30 sec. Denaturation was 98 for 10 sec, annealing was 95 for 30 sec, extension was 72 for 30 sec, step 2 to 4 was repeated 8 – 12 times depending on the DNA concentration final extension was 72 for 5 minutes.

Purification of final library

We purified the library using Beckman Coulter Agencourt AMPure XP magnetic beads (Cat no A63881). This involved mixing the cDNA with 1X of the beads and incubating the mixture, then letting it incubate on a magnetic rack and removing the supernatant. After washing the beads twice with EtOH the beads were dried and then they were eluted with Nuclease-Free water. The library was checked for quality and concentration on the Agilent 2200 TapeStation, and then frozen at -20º C until use.

Sequencing

Sequencing was performed on either the Illumina NextSeq 500 or the Ion Torrent Personal Genome Machine.

Ion Torrent

The libraries were first attached to the beads, called emulsion, and then were copied until they covered the beads, called enrichment. This was achieved either through the Ion Chef or the Ion Torrent OneTouch system. Ion Torrent technology uses semiconducting chips with millions of wells where beads with the libraries attached. The beads were flowed over the chip each deposited in a well, chip loading. The chip was placed on the sequencer and one of four nucleotides was flowed over it. When a nucleotide was incorporated into the DNA, there was a change in pH and the sequencer sensed this change and identified when that nucleotide was incorporated. Ion Torrent sequencing was performed locally at the Whitney Lab. Some libraries were also sequenced on the Ion Torrent PGM during the *Sequencing at Sea* expedition that was conducted by the lab in the spring of 2012.

Illumina

Illumina's method of sequencing is called sequencing by synthesis, because it takes a strand of DNA to be sequenced and synthesizes its complementary strand enzymatically. This method is based off of detecting the activity of DNA polymerase with a light-producing enzyme. Samples were sent to the Interdisciplinary Center for Biotechnology Research (ICBR) at the University of Florida (UF) and sequenced on the Illumina NextSeq 500 using the NextSeq 500/550 v2 Kits (Cat no. FC-404-2005).

Data Analysis

Data analysis was performed with a combination of software tools. The preliminary Ion Torrent data was downloaded directly from the PGM server onto a local computer and transferred with FileZilla (https://filezilla-project.org/) onto University of Florida's HiPerGator High Performance Computing (HPC) system, where all of the sequencing data is stored. The primary Illumina data was downloaded from basespace (https://basespace.illumina.com) when the ICBR completed sequencing. These data were downloaded onto a local computer and transferred with FileZilla onto the HPC.

CLC Genomics Workbench

Trim adapters

The files are then imported into the program CLC Genomics Workbench 7.5 (http://www.clcbio.com). The first step is to trim their adapters. The Illumina files have already had the index primers trimmed on the sequencing server as part of the workflow and these were not trimmed here. All files had Cap T30 (5'–AAGCAGTGGTATCAACGCAGAGTACT(30)) and Cap Oligo (5'-AGCAGTGGTATCAACGCAGAGTACXXXXX-3') trimmed. Ion Torrent files had Primer

A (5'- CCATCTCATCCCTGCGTGTCTCCGACTCAG – 3') and Primer B (5'-ATCACCGACTGCCCATAGAGAGGCTGAGAC - 3') trimmed as well.

Alignment

The files were aligned to the indexed genome in CLC Genomics Workbench as well. Before this was done, the genome was imported from http://research.nhgri.nih.gov/mnemiopsis/. Gff files were imported from Dr. Joe Ryan. Track files were made using the .gff file and the genome was indexed. Each project was mapped using the gene track, which included the entire genome. Mapping parameters: Mismatch cost= 2, Insertion cost= 3, Deletion cost= 3, Length fraction= 0.5, Similarity fraction = 0.8, Global alignment= no, Auto-detect paired distances= yes, Non-specific match handling= Ignore, Output mode= create stand-alone read mappings. Create report= yes. Collect un-mapped reads= No. These aligned reads were then exported as .BAM files and transferred onto the HPC.

Partek Genomics Suite

Quantification

To obtain statistical values we used Partek® Genomics Suite® software, version 6.6 Copyright ©; 2014 (46) . The *Mnemiopsis leidyi* genome was imported into Partek and annotated with the .gff file. The BAM files were imported into RNA-seq workflow and assigned the *Mnemiopsis leidyi* species and genome. Under add sample attribute > add a categorical attribute the samples are sorted into Control, 30 Min, 1 Hour and 2 Hour categories. Next we detected differentially expressed genes with mRNA quantification normalized by RPKMs with Partek by choosing mRNA quantification under the RNA-seq workflow. mRNA quantification parameters were as follows: Can assay discriminate between sense and antisence transcripts: No; In the gene-level

result report intronic reads as compatible with the gene: No; Require strict paired-end compatibility? No; Report results with no reads from any sample: No; Report unexplained regions with more than (30) reads: No. Report exon- level results: No.

One Way Analysis of Variance

Next, we performed ANOVA. We chose differential expression analysis in the RNA-seq workflow. We excluded any column that did not have a RPKM above 1 with the column filter manager. We chose to analyze the transcripts and chose the transcript_rpkm spreadsheet. We added the experimental factor 'attribute' and selected the contrast to contrast each Control vs 30 Min, Control vs 1 Hour and Control vs 2 Hour. Each of these is performed separately. These were exported into excel and most significantly differentially expressed genes were analyzed.

Next, we choose to create gene lists based off of the parameters genes which have any change in each contrast (30 vs Cont, 1Hr vs Cont, 2Hr vs Cont) with FDR (step up) < 0.05, and Fold change > 2 or Fold change < -2. With the Venn Diagram tab we took the intersection of these 3 lists to create the first gene list. We also created gene lists with a more stringent cutoff of genes that have any change (in each contrast 30 min vs Cont, 1 Hr vs Cont and 2 Hr vs Cont) with FDR (step up) < 0.005, and Fold change > 5 or Fold change < -5 which we analyzed individually.

Prediction of Secretory Peptides

As published in Moroz 2014 (11), three non-overlapping sets of results were produced, using different cutoff criteria. Multiple heuristic tools were combined to produce a secretory peptide prediction pipeline to filter and rank potential secretory products from both genome and transcriptome data. This is based off of the presence of an N-terimal signal peptide that localizes the pre-secreted peptide the endoplasmic

reticulum. This domain consists of 5-20 hydrophobic residues and is the strongest marker for secreted products. The pipeline employs the SignalP (47) and TargetP (48). These use separate Hidden Markov Model and Neural Network components, trained on a variety of organisms, for the sole purpose of secretory peptide prediction. The parameters used for creation of the secretory peptide are as follows: Use TMHMM to predict tramsmembrane domain: yes; Use Phobius to predict transmembrane domain: yes; Use Phobius to predict signal peptides: yes; SignalP score: 0.9.

Photography and Videography

The author performed all photography and videography with a Pentax K-7 Digital SLR camera. For time-lapse videography of a regenerating *Mnemiopsis*, an animal was pinned to a large petri dish coated with Silgard. Sharp forceps were used to make an epithelial injury. The recovery of this injury was recorded for 3 hours. This footage was sped up using iMovie, resulting in a 30 second video.

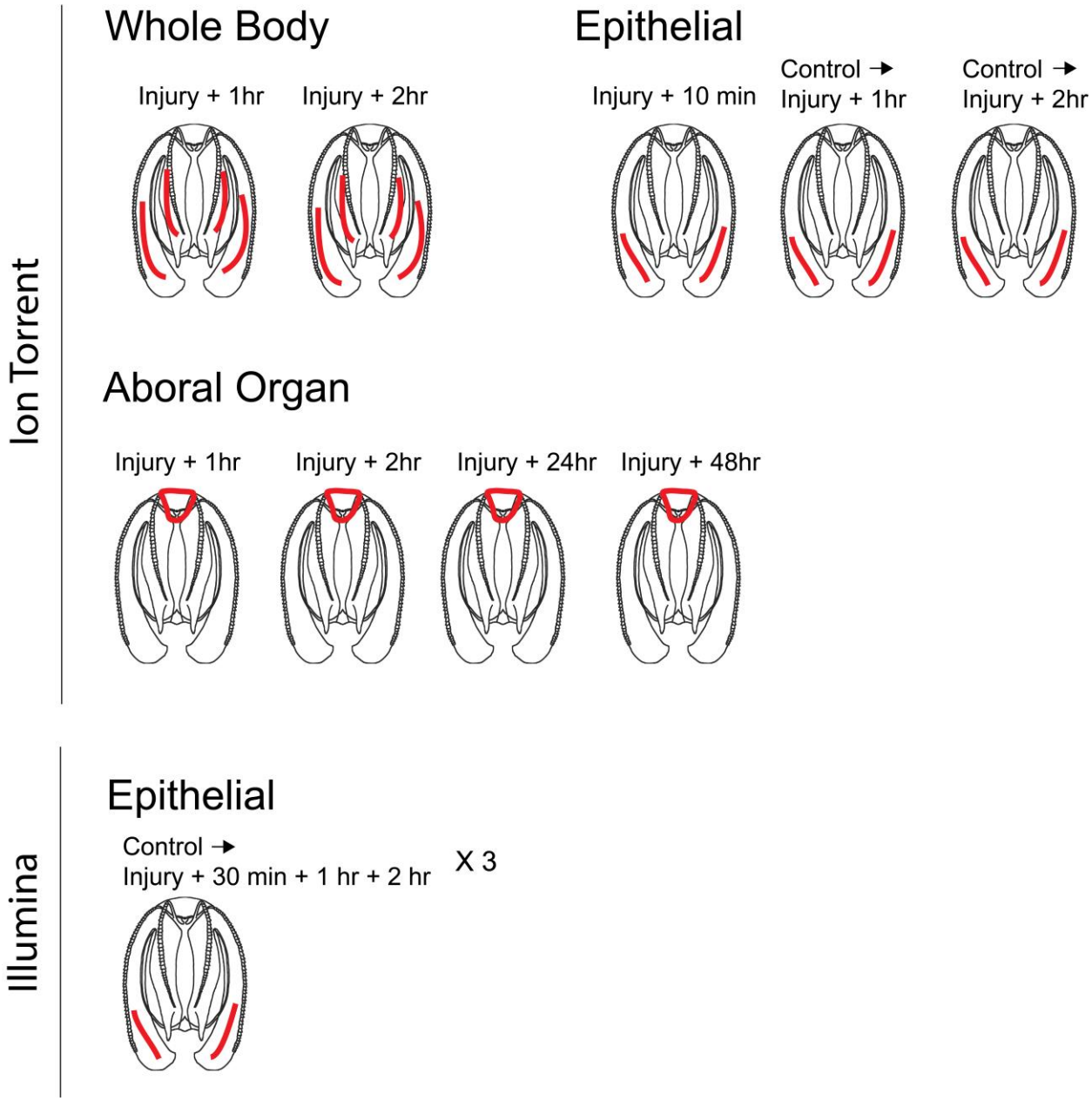


Figure 2-1. Overview of dissection procedures. Red lines indicate approximate sites of injury. Samples taken from each animal are indicated above each ctenophore. Samples used in Ion Torrent sequencing include whole body injury 1 hour, whole body injury 2 hour, epithelial injury 10 minute, epithelial control, epithelial injury 1 hour, epithelial control, epithelial injury 2 hour, aboral organ removal 1 hour, aboral organ removal 2 hour, aboral organ removal 24 hour, and aboral organ removal 48 hour. Samples used in Illumina sequencing include three replicates each of epithelial control, epithelial injury 30 min, epithelial injury 1 hour, and epithelial injury 2 hour.

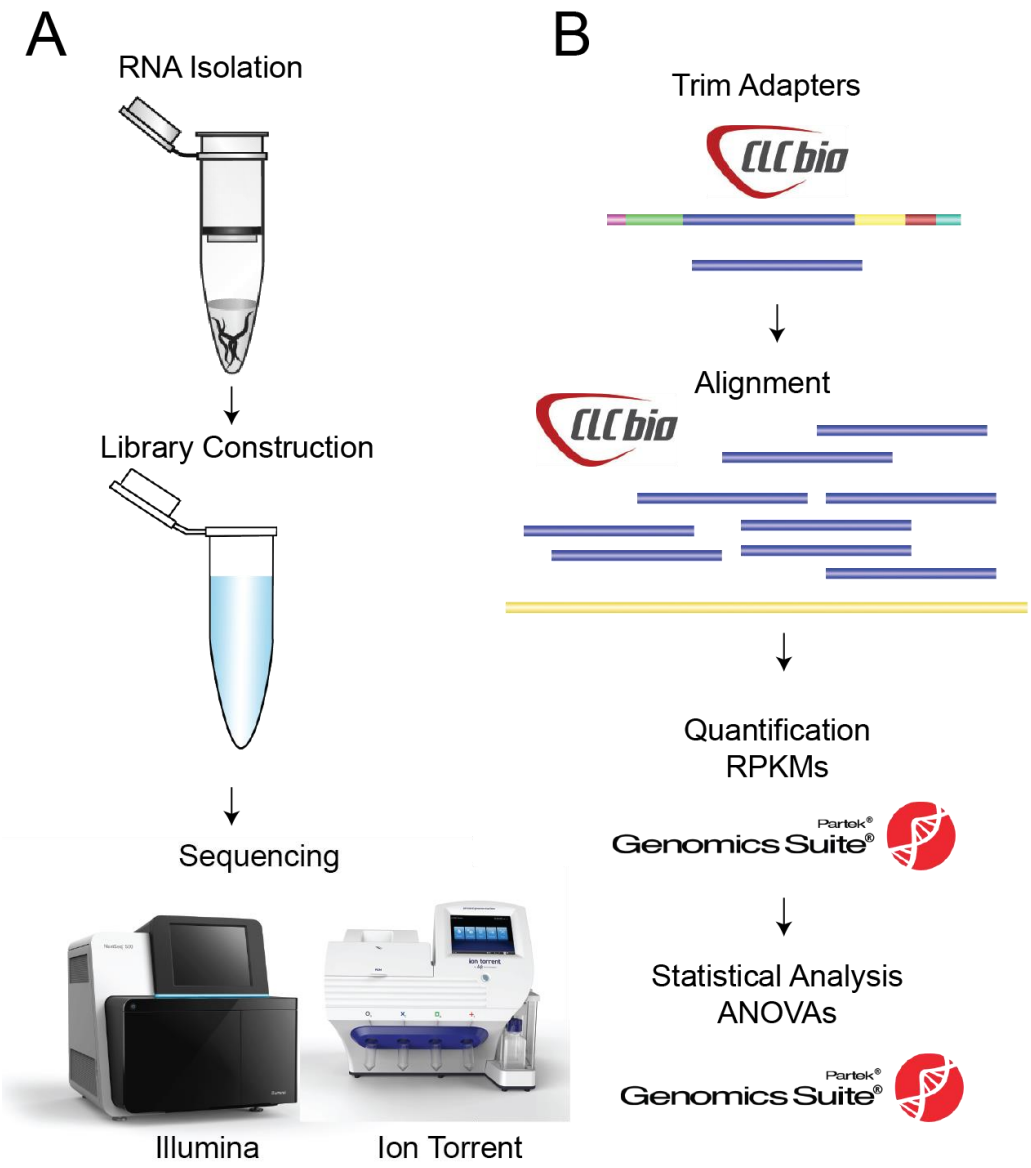


Figure 2-2. Overview of transcriptomics and data analysis methods. A) Transcriptome Sequencing. After dissection, the RNA is isolated. Next libraries are constructed. The steps of library construction are cDNA synthesis, purification of cDNA, fragmentation, end repair, adapter ligation, size selection, PCR of final library and purification of final library. Next this library is sequenced on either Illumina or Ion Torrent next generation sequencing technology. B) Data Analysis. The data from these machines are analyzed using two data analysis software CLCbio Genomic Workbench (49) and Partek Genomics Suite (46). First the adapters are trimmed in CLCbio, next the reads are aligned to the annotated genome to produce mapped .BAM files. These .BAM files are uploaded to Partek that quantifies our reads producing RPKMs or Reads Per Kilobase Per Million Mapped Reads, our measure of abundance of a gene. Last this is statically analyzed using a one way ANOVA for each gene in each project.

CHAPTER 3
RESULTS

A cumulative total of 46.1 Giga bases of sequencing data were generated. The Ion Torrent data produced 1.1 Giga bases, the RNA-Seq Illumina sequencing produced 45.1 Giga bases (Table 1). These sequencing data provide 296 X coverage of *Mnemiopsis*'s 0.1 Giga base genome.

Visual analysis of *Mnemiopsis* regeneration shows epithelial wound site closure complete within 3 hours; see Figure 3-1 and Object 3-1. See Appendix Figure A-1 for large photos.

Object 3-1 Video of Mnemiopsis Regeneration (.mp4 file 76.8 MB)

Predicted Secretory Peptides Present in Multiple Regeneration Types

Data from Ion Torrent sequencing indicate that a small list of differentially expressed genes are shared between the different tissue regeneration types, aboral organ, body injury and epithelial, and this list contains predicted secretory molecules. These data represent sequenced aboral organ regenerating tissue, epithelial regenerating tissue and whole body regenerating tissue. When these conditions were grouped by each tissue type and compared against each other, applying a cutoff of an False Discovery Rate step up (FDR) < 0.5, Fold change > 1.5 or Fold change < -1.5, there were a small group of genes which were common between the experimental groups (Fig 3-2). In this group of 33 genes 13 of these contained predicted signal peptides without transmembrane domains, which could be secretory molecules. See more information about these genes in the *Supplementary Data* provided in the Object 3-2. Here in our cut off values the fold change of -1.5 indicates a 1.5 fold decrease. For

the rest of the work a negative fold change indicates a fold decrease of that numerical value.

When a protein is destined to be a secretory protein it is produced in the ER and a sequence of amino acids tag this protein for secretion, this sequence is called the “signal peptide” in the secretory protein or secretory molecule. But not only secretory molecules have these proteins, also transmembrane proteins and proteins which end up in organelles. To predict these signal peptides a program called SignalP (50) was used. This takes the amino acid sequence of a specific transcript and predicts the location of singal peptide cleavage sites and discriminating these from transmembrane regions. An important thing to remember through this work is when discussing these secretory molecules these are all predicted via software and are not experimentally verified yet. Basically what we are seeing here that when a tissue is regenerating communication between cells is important. These predicted secretory molecules could be important in multiple regeneration types.

Object 3-2 Supplementary Data (.xls file 530KB)

Multiple Signaling Pathways Initiate Regeneration in Ctenophores

Illumina RNA-seq data provide evidence for possible involvement of multiple signal transduction pathways in regeneration of tissue in *Mnemiopsis*. These sequencing data represent three replicates each of control epithelial tissue, regenerating tissue at 30 minutes after injury, 1 hour after injury and 2 hours after injury. Any gene that did not have an RPKM above 1 in any one of these conditions was eliminated from subsequent analyses. The result was 8,121 predicted genes differentially expressed, depicted in a hierarchical clustering in Fig 3-3A. See Appendix Figure A-2A for Principal Component Analysis of this data. Replicates were grouped

and regenerating tissue compared to the control tissue. A cutoff of an FDR < 0.05, and Fold change > 2 or < -2 was applied to each experimental group, which were compared against each other resulting in the intersection of 516 genes (Fig 3-3B). These 516 genes are shown in a hierarchical clustering in Fig 3-3C, with select genes of high fold change labeled. Applying a more stringent cutoff, FDR < 0.005, Fold change > 5 or < -5, 22 genes were differentially expressed at 30 Min, 159 genes differentially expressed at 1 hour, and 132 genes at 2 hour (here after referred to as “stringent list”). Of the genes which have annotation there are 175 non redundant genes between all three 30 min, 1 hour and 2 hour stringent lists. These 175 genes are listed in Table 3-2 with their *Mnemiopsis* gene ID, Fold Change for 30 Min vs Control, 1 hour vs Control and 2 hour vs Control. See Appendix Figure A-2B for Principal Component Analysis of this stringent cut off group of data. More detailed information about these genes can be found in the *Supplemental Data*, such as the RPKM values and pfam domains. Figure 3-4 shows a hierarchical clustering of 121 currently annotated genes from the 1 hour stringent list. Of the 159 genes in the 1 hour stringent list, 38 are *Mnemiopsis* specific genes which currently have no annotation. This stringent gene list is where we searched for what the ctenophore is using to regenerate.

Secretory Peptide Signaling Pathway.

The data show robust evidence for secretory peptide signaling in *Mnemiopsis* regeneration, depicted in blue on Figure 3-4. As mentioned previously, these secretory peptides are predicted models. The stringent gene lists contain the ctenophore specific putative secretory peptide-37, originally characterized in *Pleurobrachia bachei*, and shows a 65 to 99 fold increase in regenerating tissue. These gene lists also contain two *Mnemiopsis* specific secretory peptides. MlSP29 is up-regulated 14 to 22 fold and

MlSP30 is down regulated 6 to 12 fold. Signal peptide processing genes were also present in the stringent gene lists. One of the most highly up-regulated genes for regenerating tissue in the stringent gene list is Serine/threonine-protein kinase/endoribonuclease IRE1, which is up-regulated 76 to 145 fold. Other highly up-regulated genes include the UDP-galactose translocator, Exocyst complex component 3, Furin-1, and Charged multivesicular body protein 1b. A total of 42 genes classified as signal peptide processing genes were present in the stringent gene lists between the 1 hour and 2 hour post-injury. In addition, receptors that were likely signal peptide receptors were found to be present in the stringent gene lists including Substance-K receptor, Metabotropic glutamate receptor 7, frizzled, Prolow-density lipoprotein receptor-related protein 1, and Notch1. A full list of differentially expressed genes related to signal peptide processing are provided in Table 3-2, with additional data in *Supplementary Data*. Here we show robust evidence for full involvement of the secretory peptide pathway during regeneration. All molecular components are shown to be differentially expressed including signal peptide processing proteins in the ER and the Golgi, exocytosis and endocytosis proteins, receptors for these peptides and the predicted peptides themselves.

Calcium Signaling Pathway

The data also indicate many downstream regulators of Ca^{2+}-dependent and MAP-kinase cascades to be involved in *Mnemiopsis* regeneration. These genes are indicated in red in Figure 3-4. Epidermal differentiation-specific protein, which modulates epidermal calcium, has the highest RPKM values of any gene in the stringent gene list, reaching expression levels over 6000. Also up-regulated in regeneration are Mitogen-activated protein kinase kinase kinase 10 with up regulated 5 fold, and

mitogen-activated protein kinase kinase kinase kinase 4-like with an up regulation of 15 fold. Calcium-binding mitochondrial carrier protein SCaMC-1 has a 25 to 44 fold up regulation in regeneration. Calcium related signaling has classically been involved in cells and tissues which are injured and excited and this is also shown here. These pathways seem to be evolutionarily conserved in ctenophore regeneration.

Steroid Signaling Pathway

Interestingly, two steroids were found in the stringent gene list to be up-regulated in regeneration. 3-oxo-5-alpha-steroid 4-dehydrogenase 2 was found to have a fold change of over 900,000, as it was found in the control at very low amounts. The hypothetical protein DAPPUDRAFT_305518 [Daphnia pulex], which hits to 3-oxo-5-alpha-steroid 4-dehydrogenase, is also up-regulated in regeneration. There are also three steroid binding proteins, Glutathione S-transferase 2, and Protein FAM98A (a Glutathione S-tranferase) which are down regulated in regeneration in the stringent gene list. In Figure 3-4 the steroids and steroid binding proteins are indicated in green.

Adhesion, Cytoskeleton and Energetics

The data suggest factors involved in regeneration that relate to cell adhesion, cytoskeleton rebuilding, and cell energetics. Genes involved in cell adhesion include hypothetical protein SINV_01039 [Solenopsis invicta] which contains a kringle domain and a laminin domain, as well as Interstitial collagenase B. These genes are up-regulated in regeneration in the stringent list. Cytoskeleton genes Vincullin, Kinesin-like protein KIF13A, and Dynein heavy chain-like protein PF11_0240 were also found in the stringent list to be up-regulated in regeneration. These are indicated in orange on Figure 3-4. Genes involved in cell energetics on the stringent list include ATP-citrate synthase, GTP-binding protein 10, Adenosine kinase 2, and ATP-binding cassette sub-family G

member 5. This is indicated in pink on Figure 3-4. Molecules involved in ATP binding, ATP synthesis and metabolomics which creates energy for the cell are found to be differentially expressed here in regeneration. Also, the structure of the cell, the genes involved with cytoskeleton rebuilding and adhesion are found to be up regulated as well.

Transcription Factors

In the data we found nine transcription factors in the stringent gene list. These include Carbohydrate-responsive element-binding protein, Doublesex and mab-3-related transcription factor 3, Forkhead box protein G1, LIM class homeobox transcription factor Lmx, Max-like protein X, RE1-silencing transcription factor, Transcription factor COE, Transcription factor E4F1, and Transcriptional enhancer factor TEF-3. All are up-regulated in regeneration with the exception of Max-like protein X, which is down regulated. In Figure 3-4 these are indicated in black. These transcription factors are likely up stream regulators in many of these pathways and have an important effect on regeneration.

Predicted Secretory Peptides Differentially Expressed During Regeneration

Using the signal peptide prediction pipeline available in the Moroz lab, the most stringent settings resulted in a set of 43 predicted secretory peptides. 17 of these were differentially expressed in the Illumina RNA-seq regenerating tissue. Three putative secretory peptides found in *Pleurobrachia* were added as well as the two *Mnemiopsis* specific secretory peptides found in the stringent 1 hour list. This makes a group of 21 predicted secretory peptides that are important for regeneration in *Mnemiosis* (Fig 3-5). Twelve of these peptides are up-regulated in response to injury while nine are down regulated in regeneration. MlSP3, MlSP3 and PbSP37 all have over 100 fold up regulation in regeneration, some with RPKMs over 1000. MlSP27 and PbSP31 are

down regulated genes with fold changes of below -100. We also BLASTed these secretory peptides against other ctenophore transcriptomes or gene models to see if there was any homology - these results are summarized in Figure 3-6. While these secretory peptides are predicted models the data indicates that the ctenophore may be using these as ways to communicate and organize the regeneration effort.

Early Response and Late Response Genes in Regeneration

So far the genes presented have, for the most part, either been up regulated in similar amounts all 30 min, 1 hour and 2 hour compared to control or down regulated in similar amounts 30 min, 1 hour and 2 hour. There are a smaller number of genes which were highly up regulated at 30 minutes, then less up regulated at the 1 hour and 2 hour time points, labeled early response genes. These were found by comparing the 30 min time point to the 1 hour and 2 hour, grouped, in an ANOVA test. The cutoff value here was taken from the P-value in the 30 min vs control list, as this was the most variable. I used a cut off of below 0.10. The results of this are shown in Table 3-3. There are a total of 25 early response genes, 9 containing signal peptides. The gene with the highest fold change at 30 minutes is indicated as v-Fos. This gene interacts with c-Jun, which is also highly upregulated as shown in the next section. There are also a small group of genes that were highly down regulated in 30 minutes, and less down regulated in 1 hour and 2 hour, called late response genes. These were found by the same method. These genes are shown in Table 3-4. There are a total of 12 genes, two of which have signal peptides. These early response genes will likely be signaling molecules or activators of transcription factors.

Genes Associated with Regeneration in Other Animals Generally Up-regulated in Regeneration in *Mnemiopsis*

Previous studies of regeneration have identified many gene candidates, gene families and pathways to be involved in regeneration. From eight reviews and original papers on limb regeneration(6), vertebrate regeneration(7), neuronal regeneration (9, 10, 51), zebrafish fish fin regeneration (5), muscle regeneration (52), and wound healing (8) we obtained a list of 88 genes or gene families. We found 57 of these genes to be present in the *Mnemiopsis* genome and 37 of these to be differentially expressed in regenerating tissue in the Illumina RNA-Seq data. Of those, 24 were up-regulated in regeneration and 8 were down regulated. The transcription factor c-Jun had the highest fold change, around 200 fold at 30 minutes. This gene along with c-Fos co activates transcription factor AP-1. The next most highly up regulated followed by Suppressor of Cytokine Signaling 2 and Activating Transcription Factor 6. Alpha tubulin has the highest RPKM expression value, although the fold change is approximately 1. Matrix Metalloproteinase also has a high RPKM expression value with a fold change of 11. Notch1 has the most negative fold change at -11 (Fig 3-7). Data shows that ctenophores use a number of evolutionarily conserved genes in their processes of regeneration.

Table 3-1. Number of reads, bases and giga bases present in the Illumina RNA-Seq and Ion Torrent projects.

Sequencing Technology	Reads	Bases	Gbps
Illumina RNA-Seq	305,698,456	45,052,697,944	45.05
All Ion Torrent	6,198,048	1,106,698,308	1.11
TOTAL	311,896,504	46,159,396,252	46.16

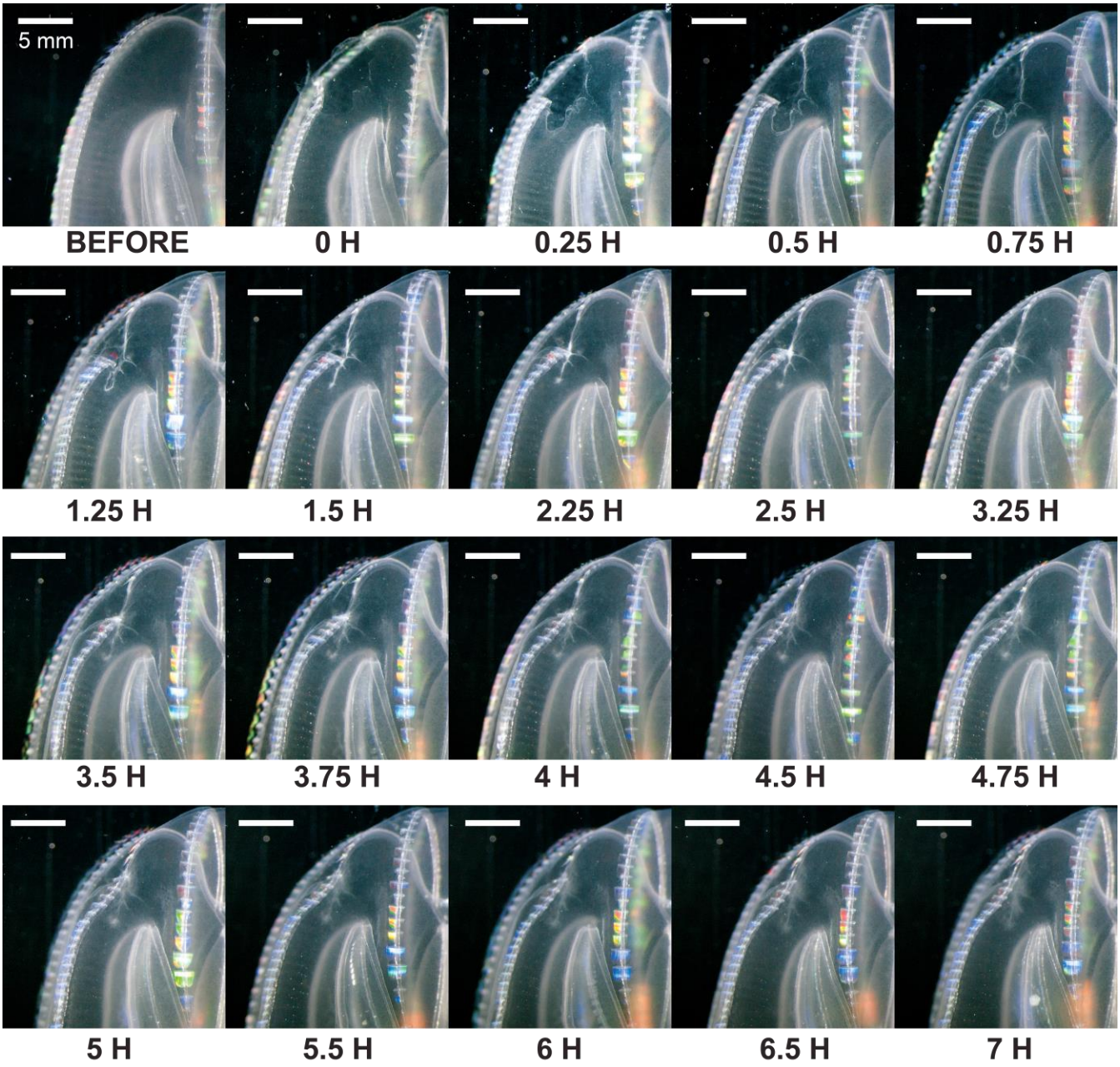


Figure 3-1. Photography of progressive *Mnemiopsis leidyi* epithelial and comb row regeneration. Total of seven hours regeneration pictured in 20 images. Epithelial wound fully closed by 3 hours as indicated by red arrow. Ctene row continues to regenerate and scar dissipates by 7 hours. Scale bar = 5 mm. Photos taken with Pentax K-7 DSLR attached to a dissecting scope with a microscope adapter. The ctenophore was pinned to a silgard dish through the procedure to reduce movement. Fresh seawater was pipetted into the dish carefully about every two hours. See Appendix Figure A-1 for larger pictures.

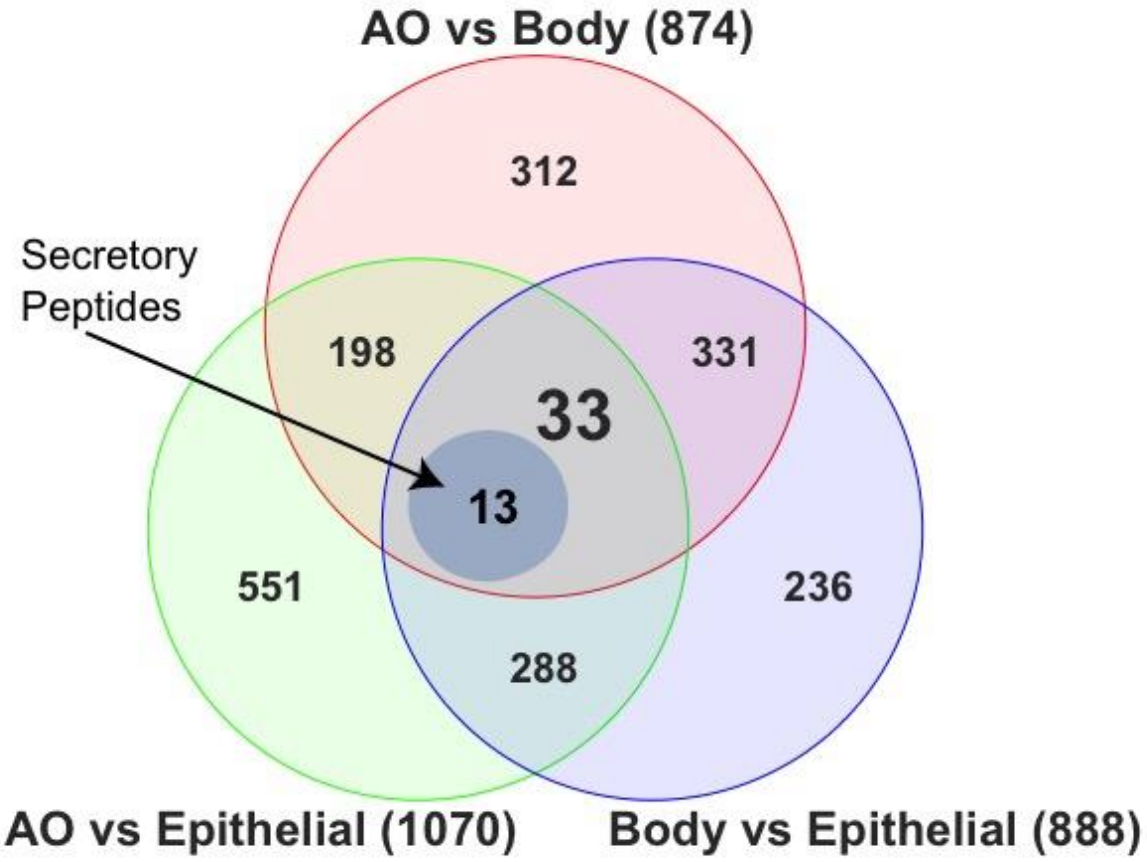


Figure 3-2. Venn diagram showing the Ion Torrent Data grouped by regeneration tissue type. These groups are then compared against each other to produce three gene lists and the intersection of these lists. The cut off for each list FDR step up < 0.5, Fold change > 1.5 or Fold change < -1.5. Intersection of each group is 33 genes, 13 of which contain predicted signal peptides in their sequences without transmembrane domains, which indicates these genes could be secretory proteins. As shown in the Illumina data, these predicted secretory proteins could play a major role in regeneration. This data shows they could also be used in other types of regeneration such as aboral organ and body regeneration. To search for the signal peptide sequence I used the website SingalP (47). The list of 33 genes is included in the Supplementary Material.

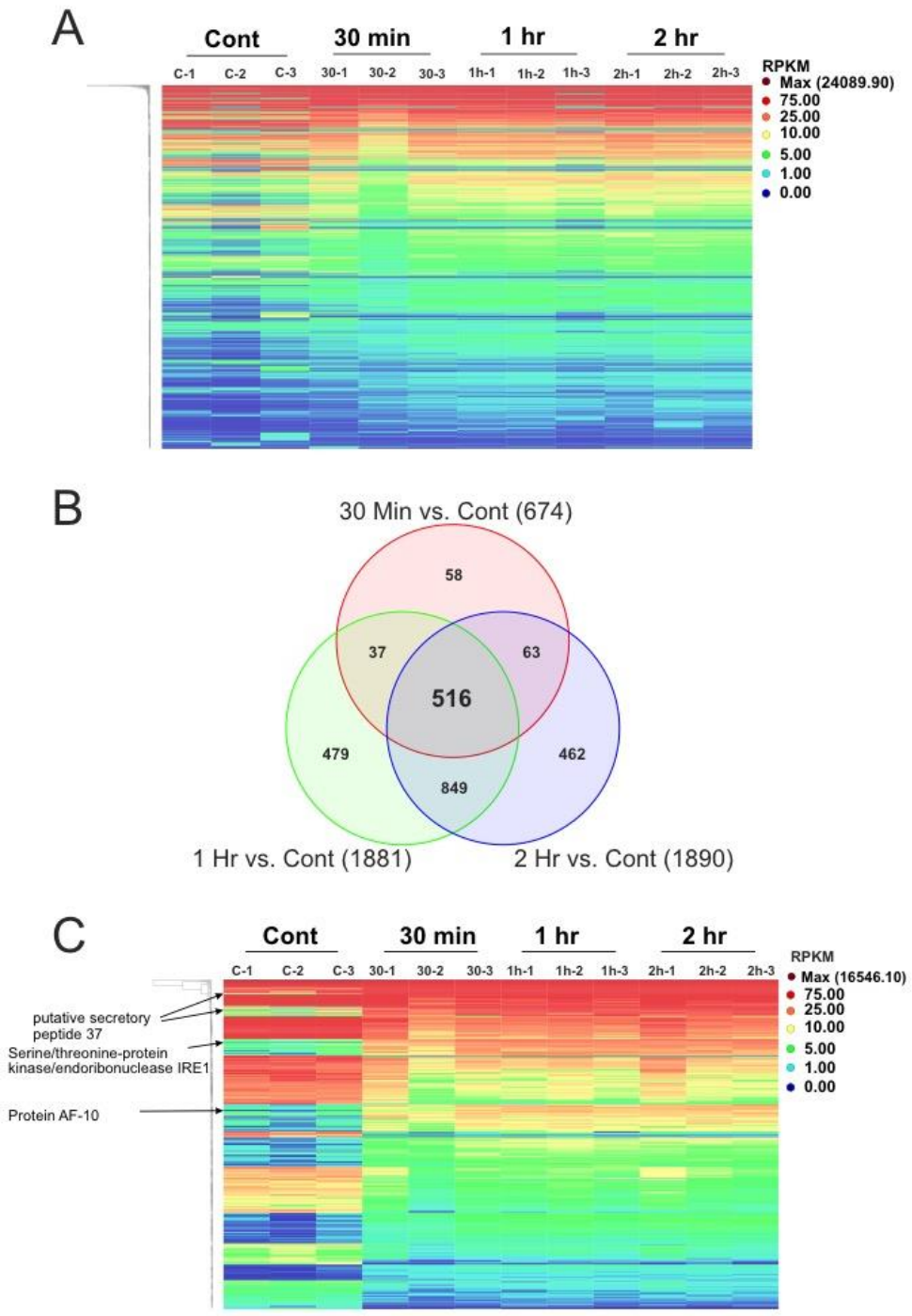


Figure 3-3. Analysis of Illumina data. A) Hierarchical clustering of all Illumina data which had an RPKM of 1 or above in any project. This totaled 8120 genes. B) First gene cut off lists. Venn diagram each gene list of FDR < 0.05, and Fold change > 2 or < -2 cut off. Intersection of these lists contains 516 genes. C) Hierarchical clustering of 516 intersection from previous gene lists. Here we need to break these down into more stringent gene lists.

Table 3-2 Stringent Gene List, as defined as FDR < 0.005, fold change > 5 or < -5

Category	Name	ID	Fold 30m v C	Fold 1h vs C	Fold 2hr vs C
Apoptotsis	poly(ADP-ribose) polymerase pme-5	ML1541114a	4.06	4.54	7.43
Apoptotsis	Cleft lip and palate transmembrane protein 1-like protein	ML225217a	4.55	8.26	10.02
Apoptotsis	Death domain-associated protein 6	ML017947a	4.87	6.56	7.62
Apoptotsis	Serine/threonine-protein kinase HT1	ML032911a	3.73	6.77	6.13
Apoptotsis	N-alpha-acetyltransferase 35, NatC auxiliary subunit	ML050825a	6.18	13.27	10.88
Ca2+ Signaling	Leucine-rich repeat protein soc-2 homolog	ML12828a	4.12	6.33	7.42
Ca2+ Signaling	BTB/POZ domain-containing protein KCTD20	ML014410a	3.67	5.56	5.32
Ca2+ Signaling	MAP kinase-activated protein kinase 2	ML310320a	7.22	15.91	20.93
Ca2+ Signaling	Histidine triad nucleotide-binding protein 1	ML458319a	-7.54	-5.97	-5.35
Ca2+ Signaling	Calcium-binding mitochondrial carrier protein SCaMC-1	ML07142a	25.93	44.46	37.34
Ca2+ Signaling	[Drosophila simulans] gb ADF78764.1	ML042718a	14.24	21.13	21.92
Ca2+ Signaling	mitogen-activated protein kinase kinase kinase kinase 4-like	ML20303a	9.77	20.51	16.67
Ca2+ Signaling	cGMP-dependent protein kinase 1	ML41322a	5.13	6.78	6.99
Ca2+ Signaling	Stromal interaction molecule 1	ML047931a	8.37	11.97	9.83
Ca2+ Signaling	Protein LURP-one-related 15	ML082713a	6.43	11.25	6.25
Ca2+ Signaling	Type II inositol 3,4-bisphosphate 4-phosphatase	ML070820a	6.34	9.30	8.21
Ca2+ Signaling	Calumenin-A	ML32781a	5.00	10.16	8.38
Ca2+ Signaling	phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit gamma	ML064510a	4.46	7.95	6.80
Ca2+ Signaling	Mitogen-activated protein kinase kinase kinase 10	ML276913a	3.55	5.93	5.13
Ca2+ Signaling	Epidermal differentiation-specific protein	ML049718a	2.05	5.27	2.97
Cell Cycle	Glycogen synthase kinase-3 beta	ML20837a	6.68	10.42	11.13
Cell Cycle	Nuclear pore complex protein Nup85	ML030416a	3.35	4.73	5.26
Cell Cycle	Leucine carboxyl methyltransferase 1	ML04636a	-8.35	-8.35	-10.61
Cell Cycle	Serrate RNA effector molecule homolog A	ML09401a	43.36	56.48	52.03
Cell Cycle	E3 ubiquitin-protein ligase ipaH3	ML11001a	14.37	27.81	33.76
Cell Cycle	DNA replication licensing factor mcm2	ML10001a	7.41	13.65	13.95
Cell Cycle	Cell cycle control protein 50A	ML12841a	5.13	7.75	7.11
Cell Cycle	E3 ubiquitin-protein ligase SMURF2	ML20687a	3.55	6.52	8.30
Cell Cycle	Ubiquitin carboxyl-terminal hydrolase 10-A	ML13565a	4.31	8.75	7.91
Cell Cycle	predicted protein [Nematostella vectensis]	ML44422a	13.02	23.40	16.70
Cell Cycle	E3 ubiquitin-protein ligase RNF170	ML049013a	6.59	11.20	9.91
Cell Cycle	E3 ubiquitin-protein ligase RNF213	ML200238a	7.95	10.44	7.90
Cell Cycle	Cullin-associated NEDD8-dissociated protein 1	ML03181a	5.04	10.04	8.05
Cell Cycle	Transforming acidic coiled-coil-containing protein 1	ML41152a	5.08	10.53	5.71
Cell Cycle	hypothetical protein LOTGIDRAFT_136165 [Lottia gigantea]	ML149613a	3.54	9.22	8.16

Table 3-2 Cont. Stringent Gene List, as defined as FDR < 0.005, fold change > 5 or < -5

Category	Name	ID	Fold 30m v C	Fold 1h vs C	Fold 2hr vs C
Cell Cycle	F-box only protein 16	ML064917a	16.46	33.41	24.37
Cell Cycle	Ubiquitin-like modifier-activating enzyme ATG7	ML220728a	4.00	7.49	5.56
Cytoskeleton	ABLIM class LIM protein	ML093519b	12.75	12.56	18.44
Cytoskeleton	hypothetical protein KGM_18974 [Danaus plexippus]	ML16038a	3.27	5.10	7.11
Cytoskeleton	Cdc42-interacting protein 4 homolog	ML03467a	3.27	4.36	5.89
Cytoskeleton	Dynactin subunit 4	ML003013a	9.62	13.59	17.21
Cytoskeleton	Kinesin-like protein KIF13A	ML020018a	10.86	15.35	17.47
Cytoskeleton	Vinculin	ML148910a	11.11	12.68	17.65
Cytoskeleton	WD repeat-containing protein 1	ML033628a	5.71	7.50	9.38
Cytoskeleton	Ataxin-3	ML04611a	-5.84	-4.36	-3.37
Cytoskeleton	Carbohydrate sulfotransferase 11-1	ML01502a	8.37	11.24	10.22
Cytoskeleton	Interstitial collagenase B	ML282519a	3.16	6.30	6.42
Cytoskeleton	Dynein heavy chain-like protein PF11_0240	ML04737a	12.63	22.50	16.41
Cytoskeleton	hypothetical protein SINV_01039 [Solenopsis invicta]	ML239527a	3.34	6.05	3.57
Detection of Pain*	Transient receptor potential cation channel subfamily A member 1	ML279614a	4.55	7.95	3.51
Energetics	MCM domain-containing protein 2	ML01385a	-4.05	-3.65	-9.24
Energetics	GTP-binding protein 10	ML005317a	-3.73	-7.26	-5.15
Energetics	proton-coupled amino acid transporter 1-like	ML073035a	16.35	25.25	19.08
Energetics	Lysosomal alpha-glucosidase	ML08481a	9.16	12.53	10.27
Energetics	ATP-citrate synthase	ML10541a	5.60	11.06	8.15
Energetics	ATP-binding cassette sub-family G member 5	ML108014a	7.27	8.61	6.72
Energetics	Adenosine kinase 2	ML00422a	3.25	7.39	5.08
Energetics	hypothetical protein CGI_10015342 [Crassostrea gigas]	ML055912a	2.81	5.01	2.98
Energetics	uncharacterized protein LOC101861715 isoform X2 [Aplysia californica]	ML35931a	17.90	34.90	34.17
Energetics	Methylmalonyl-CoA epimerase, mitochondrial	ML23999a	-3.99	-5.09	-4.74
Ion Transport	monocarboxylate transporter 9-like	ML120745a	14.84	16.88	21.93
Ion Transport	H(+)/Cl(-) exchange transporter 3	ML221319a	8.21	12.46	14.00
Ion Transport	Copper homeostasis protein cutC homolog	ML095324a	-5.89	-5.81	-6.25
Ion Transport	Protein Jade-2	ML21525a	5.66	7.64	8.42
Ion Transport	Zinc transporter ZIP1	ML21899a	12.94	23.86	16.68
Ion Transport	sodium bicarbonate transporter-like protein 11	ML063326a	3.51	6.44	4.33

Table 3-2 Cont. Stringent Gene List, as defined as FDR < 0.005, fold change > 5 or < -5

Category	Name	ID	Fold 30m v C	Fold 1h vs C	Fold 2hr vs C
Ion Transport	Carbonic anhydrase 2	ML044613a	-3.25	-5.38	-4.68
Ion Transport	zinc finger MYM-type protein 4-like	ML214012a	-5.64	-5.19	-4.48
RNA Processing	splicing factor 3B subunit 1 isoform X3	ML444213a	4.32	4.52	6.73
RNA Processing	Piwi-like protein 1	ML009119a	13.32	23.98	30.26
RNA Processing	reverse transcriptase	ML14986a	15.88	1.21	10.24
RNA Processing	RING finger protein 10	ML13372a	5.77	14.04	12.90
RNA Processing	Serine/arginine-rich splicing factor 9	ML219012a	7.39	11.36	13.04
RNA Processing	DEAD-like helicase	ML26179a	4.79	10.32	11.14
RNA Processing	SID1 transmembrane family member 1	ML11171a	4.37	5.78	9.08
RNA Processing	28S ribosomal protein S35, mitochondrial	ML084422a	4.97	5.47	5.76
RNA Processing	ELAV-like-1 [Pleurobrachia bachei]	ML111722a	3.15	6.50	5.40
RNA Processing	Transposon Ty4-H Gag-Pol polyprotein	ML00381a	6.50	11.48	7.64
RNA Processing	RNA-binding protein Nova-2	ML105427a	6.75	7.12	6.66
RNA Processing	Splicing factor U2AF 50 kDa subunit	ML02164a	3.69	8.14	5.31
Signal Peptide	putative secretory peptide-37 [Pleurobrachia bachei]	ML184414a	65.63	99.16	80.13
SP Processing	hypothetical protein LOTGIDRAFT_233561 [Lottia gigantea]	ML05232a	3.75	4.88	6.81
SP Processing	hypothetical protein PTSG_11812 [Salpingoeca rosetta]	ML35932a	5.45	6.71	9.97
SP Processing	ER degradation-enhancing alpha-mannosidase-like 1-like	ML398312a	3.70	6.00	7.82
SP Processing	Furin-1	ML279823a	3.46	4.38	5.37
SP Processing	Guanine nucleotide-binding protein subunit beta-1 sp Q61ZF6.1 GBB1	ML02234a	3.84	4.01	5.06
SP Processing	Puromycin-sensitive aminopeptidase	ML200262a	8.12	10.81	13.25
SP Processing	SCY1-like protein 2	ML010124a	18.88	32.46	36.17
SP Processing	Serine/threonine-protein kinase/endoribonuclease IRE1-2	ML47001a	-38.23	?	?
SP Processing	Putative glutathione S-transferase DHAR4	ML017940a	12.94	16.31	9.98
SP Processing	Charged multivesicular body protein 1b	ML124217a	-51.38	-15.74	-21.08
SP Processing	hypothetical protein, partial [Prochloron didemni]	ML06277a	-28.64	-51.49	?
SP Processing	Serine/threonine-protein kinase/endoribonuclease IRE1-1	ML122312a	76.82	111.84	145.52
SP Processing	Transmembrane protein 181	ML03363a	52.78	85.87	78.68
SP Processing	Protein unc-50 homolog	ML08238a	19.03	25.07	24.96
SP Processing	Serine/threonine-protein kinase/endoribonuclease IRE1	ML124218a	8.70	12.08	18.73

Table 3-2 Cont. Stringent Gene List, as defined as FDR < 0.005, fold change > 5 or < -5

Category	Name	ID	Fold 30m v C	Fold 1h vs C	Fold 2hr vs C
SP Processing	Transmembrane protein 181	ML03363a	52.78	85.87	78.68
SP Processing	Protein unc-50 homolog	ML08238a	19.03	25.07	24.96
SP Processing	Serine/threonine-protein kinase/endoribonuclease IRE1	ML124218a	8.70	12.08	18.73
SP Processing	DENN domain-containing protein 5A	ML03452a	6.79	8.39	9.75
SP Processing	Exocyst complex component 3	ML02876a	5.19	8.67	8.45
SP Processing	Polypeptide N-acetylgalactosaminyltransferase 1	ML056920a	4.69	7.99	8.24
SP Processing	Phosphatidylinositol-binding clathrin assembly protein LAP	ML06092a	4.54	7.69	7.27
SP Processing	predicted protein [Nematostella vectensis] gb EDO46646.1	ML30075a	3.82	6.31	6.19
SP Processing	predicted protein [Nematostella vectensis] gb EDO47167.1	ML28207a	3.41	5.81	6.36
SP Processing	Nicastrin	ML102219a	3.04	5.10	5.41
SP Processing	predicted protein [Nematostella vectensis]	ML25826a	6.48	12.81	10.51
SP Processing	Carbohydrate sulfotransferase 11	ML010122a	4.25	8.25	7.61
SP Processing	Epsin-2	ML03583a	11.30	20.17	13.42
SP Processing	Reticulon-4	ML204414a	9.33	18.69	14.99
SP Processing	Protein phosphatase 1 regulatory subunit 37	ML21881a	5.92	12.97	9.32
SP Processing	UDP-galactose translocator	ML09354a	7.28	12.40	8.36
SP Processing	ERO1-like protein alpha	ML154141a	4.85	10.48	7.19
SP Processing	Endoplasmic reticulum aminopeptidase 2	ML030513a	5.21	8.90	8.36
SP Processing	Transmembrane and coiled-coil domain-containing protein 4	ML11271a	4.76	7.82	6.67
SP Processing	Solute carrier family 35 member F5	ML190418a	4.73	7.40	6.13
SP Processing	Vacuolar protein sorting-associated protein 8 homolog	ML02314a	4.27	6.95	6.79
SP Processing	CAAX prenyl protease 1 homolog	ML007427a	3.78	6.51	4.96
SP Processing	Trafficking kinesin-binding protein 1	ML074227a	4.28	5.15	4.84
SP Processing	AP-4 complex subunit epsilon-1	ML002138a	4.27	5.65	3.67
SP Processing	Major facilitator superfamily domain-containing protein 1	ML123613a	3.58	5.53	3.73
SP Processing	Signal recognition particle 19 kDa protein	ML06834a	-3.83	-6.67	-4.05
SP Processing	Peptidyl-prolyl cis-trans isomerase	ML43115a	-3.60	-7.54	-5.04
SP Processing	WD repeat domain phosphoinositide-interacting protein 4	ML342215a	4.17	6.53	5.58
SP Processing	Leucine-zipper-like transcriptional regulator 1	ML22139a	3.29	6.07	5.13
SP Processing	predicted protein [Nematostella vectensis] gb EDO36671.1	ML07245a	2.95	5.10	4.92
SP Receptor	Prolow-density lipoprotein receptor-related protein 1	ML223521a	2.36	3.37	5.63

Table 3-2 Cont. Stringent Gene List, as defined as FDR < 0.005, fold change > 5 or < -5

Category	Name	ID	Fold 30m v C	Fold 1h vs C	Fold 2hr vs C
SP Receptor	neurogenic locus notch homolog protein 1-like	ML199821a	-9.00	-11.02	-19.65
SP Receptor	frizzled [Mnemiopsis leidyi]	ML003224a	5.85	8.40	8.57
SP Receptor	2-oxoglutarate receptor 1, partial	ML310311a	18.53	124.42	64.12
SP Receptor	hypothetical protein AURANDRAFT_62439 [Aureococcus anophagefferens]	ML35871a	32.03	72.42	31.43
SP Receptor	Substance-K receptor	ML096818a	3.16	7.08	4.20
SP Receptor	Metabotropic glutamate receptor 7	ML071311a	2.21	6.52	3.87
Steriod Binding	estrogen sulfotransferase-like	ML04646a	2.45	4.03	8.00
Steroid	hypothetical protein DAPPUDRAFT_305518 [Daphnia pulex] (3-oxo-5-alpha-steroid 4-dehydrogenase)	ML008120a	5.01	6.46	6.82
Steroid	3-oxo-5-alpha-steroid 4-dehydrogenase 2	ML05085a	266828.0 0	1955350.0 0	484490.0 0
Steroid Binding	Glutathione S-transferase 2	ML01926a	-11.04	-5.97	-4.88
Transcription Factor	transcription factor COE [Mnemiopsis leidyi]	ML04474a	4.97	6.93	9.19
Transcription Factor	Doublesex- and mab-3-related transcription factor 3	ML008118a	6.69	11.50	13.90
Transcription Factor	Transcription factor E4F1	ML03178a	2.16	1.69	6.01
Transcription Factor	Max-like protein X	ML20685a	-8.95	-3.90	-4.23
Transcription Factor	Transcriptional enhancer factor TEF-3	ML200210a	23.47	23.63	12.84
Transcription Factor	LIM class homeobox transcription factor Lmx	ML11549a	17.34	21.68	8.22
Transcription Factor	Forkhead box protein G1	ML154122a	10.64	20.91	13.97
Transcription Factor	Carbohydrate-responsive element-binding protein	ML070810a	6.98	9.10	7.85
Transcription Factor	RE1-silencing transcription factor	ML086428a	3.35	7.81	5.90
Other	predicted protein [Micromonas sp. RCC299]	ML015415a	10.11	13.01	13.71
Other	predicted protein [Nematostella vectensis]	ML00087a	3.52	5.58	5.95
Other	U-box domain containing protein [Tetrahymena thermophila]	ML070218a	6.86	11.09	15.18
Other	Beta-1,3-galactosyltransferase 5	ML06709a	8.11	5.53	13.46
Other	Bromodomain testis-specific protein	ML073037a	3.16	4.37	5.42
Other	Cholesterol 24-hydroxylase	ML19912a	7.71	11.07	14.09
Other	Putative ATP-dependent RNA helicase DHX33	ML282530a	2.86	9.20	13.99
Other	Ethanolamine kinase 1	ML17034a	7.97	9.90	15.23
Other	Far upstream element-binding protein 1	ML14308a	4.90	4.95	6.73
Other	Glycosyltransferase 8 domain-containing protein 1	ML36898a	8.50	13.61	15.40
Other	Myotubularin-related protein 10-B	ML034661a	4.89	7.80	9.09
Other	Probable NADH dehydrogenase	ML42441a	14.75	23.49	26.71

Table 3-2 Cont. Stringent Gene List, as defined as FDR < 0.005, fold change > 5 or < -5

Category	Name	ID	Fold 30m v C	Fold 1h vs C	Fold 2hr vs C
Other	Epidermal retinol dehydrogenase 2	ML094321a	4.09	7.11	10.91
Other	Phosphoserine aminotransferase	ML08922a	3.78	6.37	7.45
Other	Putative transporter SVOPL	ML25281a	5.89	6.70	8.32
Other	TBC1 domain family member 10A	ML073220a	4.44	5.99	7.29
Other	Thioredoxin-related transmembrane protein 1	ML154129a	3.11	5.82	7.10
Other	Zinc finger protein VAR3, chloroplastic	ML199110a	4.81	4.42	9.27
Other	Uncharacterized protein L116 [Acanthamoeba polyphaga mimivirus]	ML043314a	7.75	9.41	12.19
Other	uncharacterized protein LOC105557001, partial [Vollenhovia emeryi]	ML31986a	5.78	1.63	3.28
Other	hypothetical protein Tcan_00622, partial [Toxocara canis]	ML06152a	4.37	6.93	6.95
Other	Hypothetical protein CBG05275 [Caenorhabditis briggsae]	ML09148a	2.47	6.98	6.32
Other	DNA-directed RNA polymerase I subunit RPA12	ML005337a	-4.62	-6.24	-5.12
Other	hypothetical protein PHYSODRAFT_336759 [Phytophthora sojae]	ML43781a	9.58	16.78	9.87
Other	Ankyrin repeat domain-containing protein SOWAHC	ML15044a	6.10	8.45	6.75
Other	Proto-oncogene c-Fos	ML09961a	4.72	7.95	4.26
Other	predicted protein [Nematostella vectensis] gb EDO42914.1	ML234511a	2.24	5.63	3.94
Other	Reactive oxygen species modulator 1	ML15414a	-4.20	-6.58	-4.20
Other	hypothetical protein PPTG_03615 [Phytophthora parasitica]	ML16923a	-3.39	-15.03	-4.32

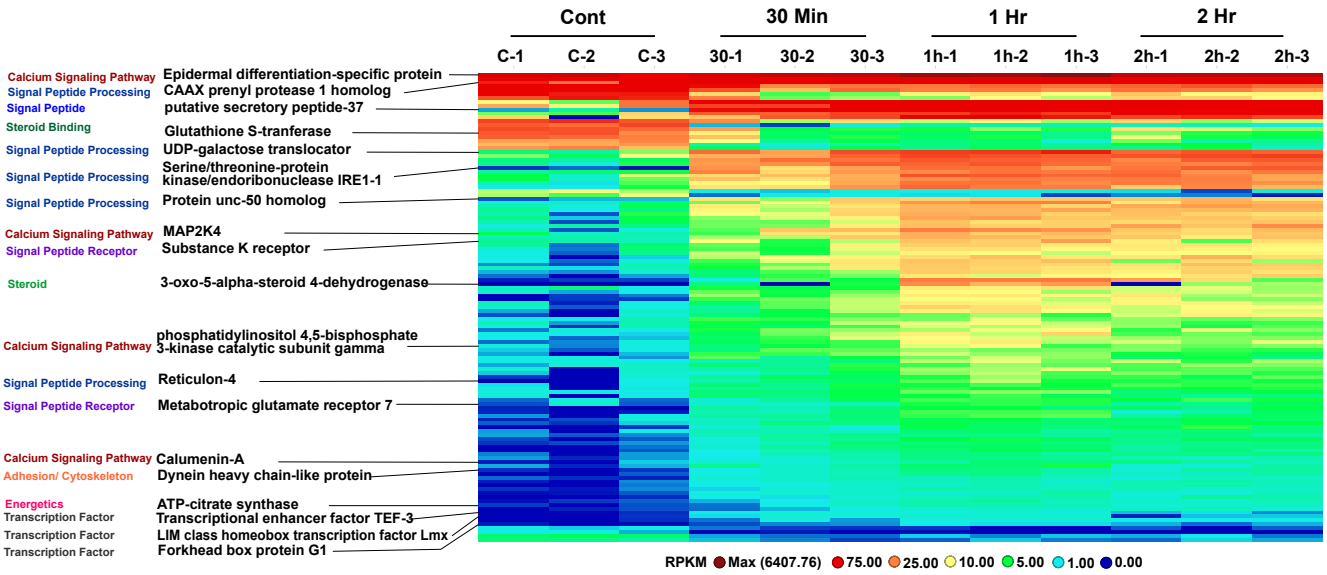


Figure 3-4. Hierarchical clustering of Illumina data stringent subset data as defined as FDR < 0.005, fold change > 5 or < -5 for the 1 Hr vs Control gene list. This hierarchical clustering contains 121 genes which represent the stringent gene list of data that was analyzed to find the processes in ctenophore tissue that are essential in regeneration. We found these to include a signal peptide pathway, calcium signaling pathway, steroid signaling, energetic molecules, adhesion molecules, transcription factors and more. Some of these genes are identified on this hierarchical clustering and the category they relate to.

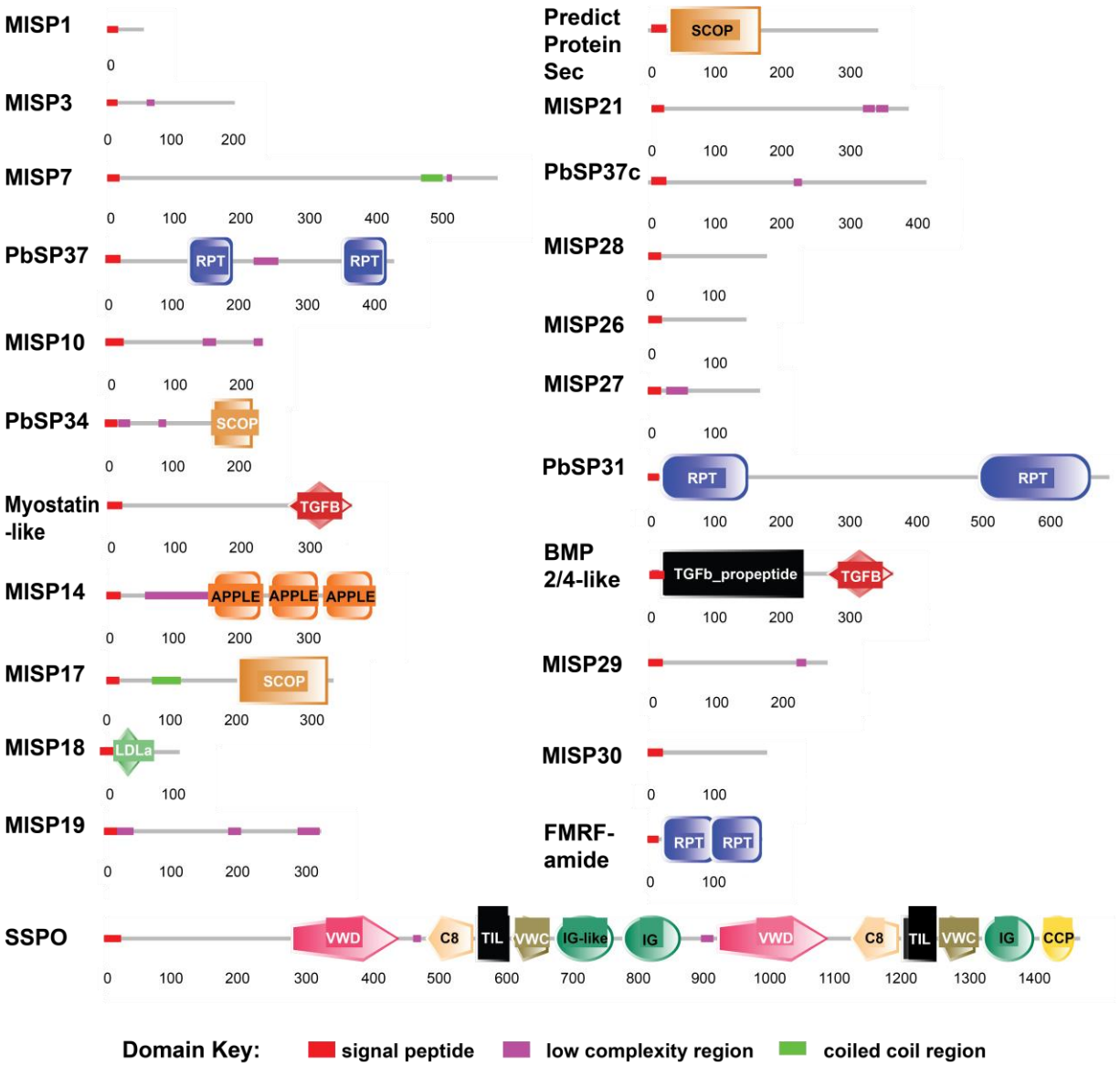


Figure 3-5. Predicted secretory peptides shown by their domains. Forty three predicted secretory molecules were predicted using the secretory prediction pipeline, which predicts secretory molecules based on presense of a signal peptide and absence of a transmembrane domain. Out of these predicted secretory peptides. Seventeen of the 43 were differentially expressed in regeneration tissue. These plus five from the stringent gene list makes 22 predicted secretory molecules investigated. This diagram is the predicted secretory molecules as shown by their domains and amino acid length indicated. Images modified from SMART (53) domain images. Red block = signal peptide, purple block = low complexity region, green block = coiled coil region. RPT = repeat domain, SCOP = structural classification of proteins, TGFB= Transforming growth factor-beta (TGF-beta) family, APPLE = Apple domain, LDLa = Low-density lipoprotein receptor domain class A, VWD = von Willebrand factor (vWF) type D domain, C8 = 8 conserved cysteine residues, TIL = Trypsin Inhibitor like cysteine rich domain, VWC = von Willebrand factor (vWF) type C domain, IG-like = Immunoglobulin like, IG = Immunoglobulin, CCP = complement control proteins.

Gene	*Pleurobrachia*	*Cestum*	*Antartic Beroe*	*Coeloplana*	*Ocyropsis*	*Lobatoplampa*
MlSP28	✓	✓	✓	✓	✓	✓
Myostatin-like	✓	-	-	✓	✓	-
PredictProteinSec	✓	✓	-	✓	✓	✓
PbSP31	✓	✓	-	✓	✓	✓
PbSP34	✓	✓	✓	✓	✓	✓
PbSP37	✓	✓	✓	✓	✓	✓
PbSP37c	✓	✓	✓	✓	✓	✓
SSPO	✓	✓	✓	✓	✓	✓
MlSP26	-	✓	-	-	✓	✓
MlSP1	-	✓	✓	✓	-	✓
MlSP10	-	✓	✓	-	✓	-
MlSP14	✓	✓	✓	✓	✓	✓
MlSP17	✓	✓	✓	✓	✓	✓
MlSP18	-	✓	-	-	-	-
MlSP19	✓	✓	✓	-	✓	-
MlSP21	-	✓	-	-	✓	-
MlSP27	-	✓	-	-	-	-
MlSP29	-	-	-	-	-	-
MlSP30	-	✓	-	-	✓	-
MlSP3	-	✓	-	-	✓	✓
MlSP7	✓	✓	✓	✓	✓	✓
BMP2/4-like	✓	✓	✓	✓	✓	✓

Figure 3-6. Predicted secretory peptide homology to other ctenophore species. BLAST against each transcriptome (or gene model for *Pleurobrachia*) was performed and if there was a match with an e-value of below 1e-5 this was considered homologous and a green box with a check is shown. Table with specific ID numbers and e values is provided in the *Supplemental Data*. Some of these secretory molecules are conserved across ctenophore species, such as PbSP37c, but others like MlSP29 seem to only be present in *Mnemiopsis*. Though, because these are transcriptomes of certain states and not genomes or gene models for most of these organisms it is not definitive that these molecules were just not expressed at this certain condition. These transcriptomes and gene models can be accessed on http://neurobase.rc.ufl.edu/pleurobrachia/browse.

Table 3-3. Early response genes. These genes are highly up regulated at 30 Min, then less up regulated at 1 Hr and 2 Hr, or slightly down regulated.

Name	ID	Fold 30m v C	Fold 1hr v C	Fold 2h v C
Transforming protein v-Fos/v-Fox	ML182032a	158.46	30.29	15.24
No annotation*	ML003232a	142.24	47.28	8.19
C-type mannose receptor 2-like*	ML106622a	105.27	16.21	31.92
No annotation	ML006310a	35.11	14.67	2.53
PHD finger protein ALFIN-LIKE 2	ML218810a	33.87	7.24	10.50
Hemicentin-1*	ML078919a	33.05	7.80	3.33
Dynein heavy chain 7	ML046520a	25.51	5.89	9.10
homeobox transcription factor HD70	ML26871a	23.34	2.60	11.62
Arylsulfatase B*	ML074251a	21.76	-1.10	10.32
No annotation*	ML16906a	18.04	6.24	1.80
predicted protein gb EDO33301.1 [Nematostella vectensis]	ML17892a	16.85	1.54	6.32
hypothetical protein KNP414_06052*	ML043317a	16.17	1.88	7.01
No annotation	ML104347a	16.01	4.56	2.88
homeobox transcription factor HD70	ML01431a	15.31	1.16	6.99
MICAL class LIM protein ML223524b*	ML05096a	14.02	-1.19	4.08
No annotation	ML14711a	12.04	2.61	4.04
GAS2-like protein 3	ML075212a	11.79	1.52	1.25
No annotation	ML141128a	4.72	1.47	-1.96
No annotation*	ML087114a	4.10	1.09	-4.04
No annotation	ML13836a	3.84	-12.79	1.73
No annotation	ML08305a	3.55	-1.57	1.29
No annotation	ML327418a	3.35	-1.15	-1.14
Neurofilament heavy polypeptide	ML017926a	3.27	-4.73	1.38
No annotation*	ML071165a	3.12	-1.61	-1.10
hypothetical protein, partial	ML15098a	2.64	-1.32	-4.53

*Gene contains a signal peptide.

Table 3-4 Late response genes. These genes are highly down regulated at 30 Min, but much less down regulated at 1 Hr and 2 Hr.

Name	ID	Fold 30m v C	Fold 1h vs C	Fold 2h vs C
glutathione S-transferase	ML343422a	-326.87	-5.27	-54.98
No annotation	ML343421a	-266.93	-3.63	-66.58
Glutathione S-transferase	ML266619a	-238.15	-3.26	-37.56
No annotation	ML305527a	-122.50	-1.56	-62.51
photoprotein 9	ML34231a	-81.88	1.02	-17.48
No annotation*	ML10665a	-70.49	-5.47	-6.82
Pentraxin-4*	ML08033a	-26.11	-1.67	-1.86
uncharacterized protein LOC105319638 [Crassostrea gigas]	ML009816a	-19.62	-3.97	-4.87
No annotation	ML13551a	-13.91	-2.96	-2.90
mucin-like protein	ML009817a	-8.04	1.38	-1.49
No annotation	ML08577a	-7.41	-1.24	-2.43
No annotation	ML051416a	-5.97	1.19	-2.78

*Gene contains a signal peptide.

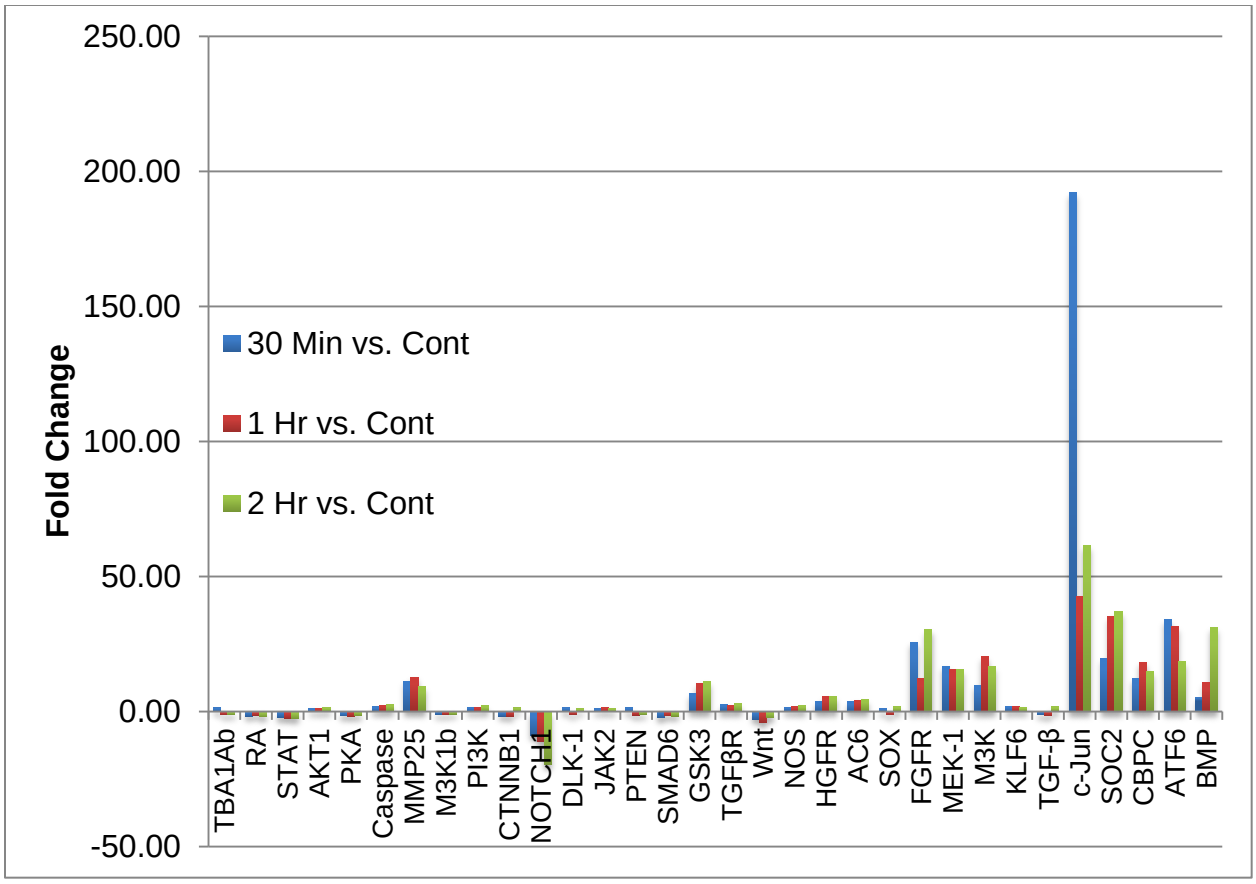


Figure 3-7. Fold change of the genes which have previously been associated with regeneration and also show differential expression in *Mnemiopsis*. From studies on limb regeneration (6), vertebrate regeneration (7), neuronal regeneration (9, 10), zebrafish fin regeneration (5), muscle regeneration (52) and wound healing (8) we obtained a list of 88 genes, 56 of which *Mnemiospsis* contained in its gene models. Thirty-two were differentially expressed (shown here), 24 up regulated and 8 down regulated. The gene with the highest positive fold change is c-Jun, which is very highly up regulated at 30 minutes having almost a 200 fold change. This gene might interact with v-Fos from the Early Response genes which also has a high up regulation at 30 minutes, forming the AP-1 early response transcription factor. Genes which are down regulated in this list are Notch, SMAD6, Wnt, RA, STAT, CTNNB1, and TGF-B, which have a negative fold change or a down regulation in regeneration. Additional information on these genes is present in the *Supplementary Data*.

CHAPTER 4
DISCUSSION AND CONCLUSION

Predicted Secretory Peptides Present in Multiple Regeneration Types

The Ion Torrent data indicate that ctenophores employ secretory peptides, as well as other key genes, in all types of tissue regeneration. The secretory peptides Spondin, which is secreted by cells in the floor plate and is involved in neural cell adhesion (54), putative secretory peptide-31 and the *Mnemiopsis* specific MlSP14 are shared between these regeneration types and these also appear in the Illumina RNA-Seq differentially expressed data. As with any limited data set, more sequencing will need to be done on the different regeneration types, such as the aboral organ and body injury, to confirm the findings. Nonetheless, the high RPKM values for these secretory peptides in multiple regenerating tissue types provide valuable evidence for their importance across multiple regeneration regimens.

Multiple Signaling Pathways Initiate Regeneration in Ctenophores

Illumina RNA-Seq data suggest that ctenophores employ a unique combination of evolutionarily conserved signal transduction pathways and ctenophore specific signaling molecules during regeneration. These animals integrate conserved proteins and ctenophore innovations into a signal peptide pathway, as well as employing evolutionarily conserved Ca^{2+}-signaling and steroid signaling pathways. The multiple signaling pathways, as well as multiple signal peptides suggests that ctenophores use parallel pathways to achieve similar biological outcomes. This system follows the established standard that degeneracy, the ability of multiple elements that are structurally different to perform the same function, is the key to a well-functioning system (55). Within the stringent gene list, which is undoubtedly excluding some genes

that are involved, multiple components of each signal transduction pathway are indicated. The signal transduction pathway has all components present, the signal peptide, 42 processing proteins, and receptors. The Ca^{2+}-pathway has multiple calcium binding proteins and multiple MAP-kinases. The steroid pathway has both steroids and steroid binding proteins.

In the calcium-signaling pathway some of the genes found to be differentially expressed in these pathways include Calnumerin-A, which is a Ca^{2+} binding protein but also has a role in the ER (56). Histidine triad nucleotide-binding protein 1 (HINT1) exhibits tumor suppressor properties and regulates Ca^{2+} signaling in mouse fibroblasts (57). Epidermal differentiation-specific protein is a Beta/Gamma crystallin protein which contain a universal motif for binding calcium (58). This gene has one of the highest RPKM values in our stringent gene list and is up regulated with regeneration. In the secretory signaling pathway some of the genes found to be differentially expressed with regeneration include Furin-1, a protease that activates large numbers of proprotein substrates and is important in apoptosis (59). Serine/threonine-protein kinase/endoribonuclease IRE1 promote cell survival during ER stress and the unfolded protein response by reducing misfolded protein levels (60). One of the transcripts of IRE1 is the highest fold change in the stringent gene list, around 100 fold change in the regeneration response. Another highly up regulated gene, Transmembrane protein 181 mediates the action of a class of toxins which display DNase activity and disrupt the cell cycle in human cell lines (61).

We also found nine transcription factors in the stringent gene list. Four of these transcription factors have had past studies that associated them with regeneration

including Forkhead box (62), RE1-silencing transcription factor (63, 64), transcription factor COE (65), and LIM class homeobox transcription factor Lmx (66). Other transcription factors that were differentially expressed included Carbohydrate-responsive element-binding protein and its binding partner Max-like protein X which classically regulate glucose and lipid metabolism (67) and could be controlling the energetics of the regeneration process. Also expressed are Transcription factor E4F1, shown to be associated with cell cycle function(68), Transcriptional enhancer factor TEF-3 associated with cell proliferation (69), and Doublesex- and mab-3-related transcription factor 3. These transcription factors are likely upstream regulators of the signaling pathways we have outlined. They could be playing similar roles as previous studies indicate, or ctenophores could be adapting these genes to these signal transduction pathways to better coordinate the regeneration process.

This stringent gene list has been narrowed down from 8,121 genes differentially expressed and 46.4 Gb of sequencing data. Each of these proteins shows highly significant differential gene expression in regenerative tissue. These transcription factors, signal molecules, steroids, processing proteins, receptors, energetic molecules, adhesion factors and cytoskeleton components combine to regulate regeneration in ctenophores as depicted in Figure 4-1.

Secretory Peptides Differentially Expressed During Regeneration

Secretory peptides are proteins that N-terminal contains a short signal sequence of 5 to 30 amino acids of the newly synthesized protein that directs the protein from the cytosol to a target site (70). Using the predictive software and our most stringent cut off, a SignalP score of 0.9, in our predicted secretory peptide pipeline we found 42 that met this criteria. 28 of these predicted secretory peptides are *Mnemiopsis* specific, and four

more ctenophore specific. Almost half, sixteen, of these peptides were differentially expressed during regeneration and the majority of these, thirteen, are ctenophore specific secretory molecules. Additionally we found two more *Mnemiopsis* specific secretory molecules and one more ctenophore specific secretory molecule in our stringent gene list. In a phylogenic tree made by RaxML (See Appendix Figure A-3) all of these secretory peptides had very low support for clustering together. This suggests that these could be all new, derived genes instead of being conserved. From this data we can conclude that ctenophores could be producing a number of unique secretory peptides that they employ during regeneration and these secretory peptides could be the key as to how ctenophores regenerate so rapidly. Here, we do have to keep in mind that these secretory peptides are predicted models need to be experimentally verified.

Early Response and Late Response Genes in Regeneration

The genes that have a much higher fold change at 30 minutes, than at 1 hour or 2 hours are suspected to be genes that are going to be the early signaling molecules in regeneration. This is supported by the result that 9 out of 25 of the genes found here have signal peptides present. An important gene found here was v-Fos, an early response gene found to be highly upregulated at 30 minutes. This gene is similar to a gene which interacts with c-Fos which was found in the regeneration associated genes in the next section.

Many Regeneration Associated Genes are Up-regulated in Regeneration in *Mnemiopsis*

A broad range of regeneration types was surveyed from zebrafish fin regeneration (5) to wound healing (8) to gather a wide array of genes to analyze. The result that out of 88 genes or gene families that were looked at, 57 were present in the

Mnemiopsis genome tell us even though ctenophores are innovative in many of their mechanisms they also have a lot of conservative genes. Additionally, the result that 32 of the previously studied regeneration associated genes were differentially expressed in injury indicates similarity between ctenophore regeneration and regeneration in the rest of the Metazoa. In Figure 4-2 the regeneration genes are broken down by the associated regeneration categories and the *Mnemiopsis* expression. The gene with the highest fold change from this group is c-Jun. This gene is up regulated almost 200 fold at 30 minutes. C-Jun interacts with c-Fos, a similar gene v-Fos was found to be in the early response genes also very highly up regulated at 30 minutes. This complex forms the AP-1 transcription factor. This is a activator protein which is a master regulator of gene expression to a variety of cellular processes such as proliferation, apoptosis and differentiation. This up regulation of these genes likely controls much activity after regeneration. Additionally, the neuronal regeneration set has the most genes that have an increase in regeneration in *Mnemiopsis*, which might be due to the fact that this category has almost double the amount of genes as the others, but could also be because nerve net covers the epithelium of ctenophores and these genes regulate the regeneration of this tissue.

Conclusion

Investigations of the molecular components and mechanisms behind regeneration are a growing field. This is the first study to analyze the molecular and genetic components behind regeneration in the phylum Ctenophora, which are known for their regenerative capabilities. Findings from these investigations provide essential insights into this phylum and regeneration machinery in an evolutionarily derived taxa.

Cumulatively, the data show that ctenophores efficiently employ a combination of evolutionarily conserved signal transduction pathways and unique ctenophore specific secretory peptides to achieve regeneration. These animals use evolutionarily conserved pathways and transcription factors that have been associated with regeneration, and some that have not previously been shown to be associated with regeneration. Even with extremely stringent data settings, multiple signaling pathways are shown to be active and at least 3 different ctenophore specific signal peptides differentially expressed. This suggests these animals are using multiple signaling pathways in parallel to achieve regeneration, demonstrating the concept of degeneracy (55). This approach provides flexibility in the processes which mediate regeneration and could explain why these animals demonstrate such proficient regenerative capabilities. Ctenophore specific secretory molecules also play a large role in the successful regeneration of these animals. These aptly named molecules are generally secreted, and are part of intracellular signaling or cell-cell signaling interactions which are critical after a cell or tissue needs to regenerate after injury. Again, with very stringent criteria, we have found 16 predicted secretory molecules that are specific to ctenophores. These multiple secretory molecules also support the idea of parallel or degenerate pathways to regeneration in this animal. In conclusion, ctenophores employ a unique set of molecular mechanisms including evolutionarily conserved transduction pathways, ctenophore specific secretory molecules, and regeneration related conserved genes all with multiple pathways to be able to be flexible in their mechanism regeneration to achieve desired results.

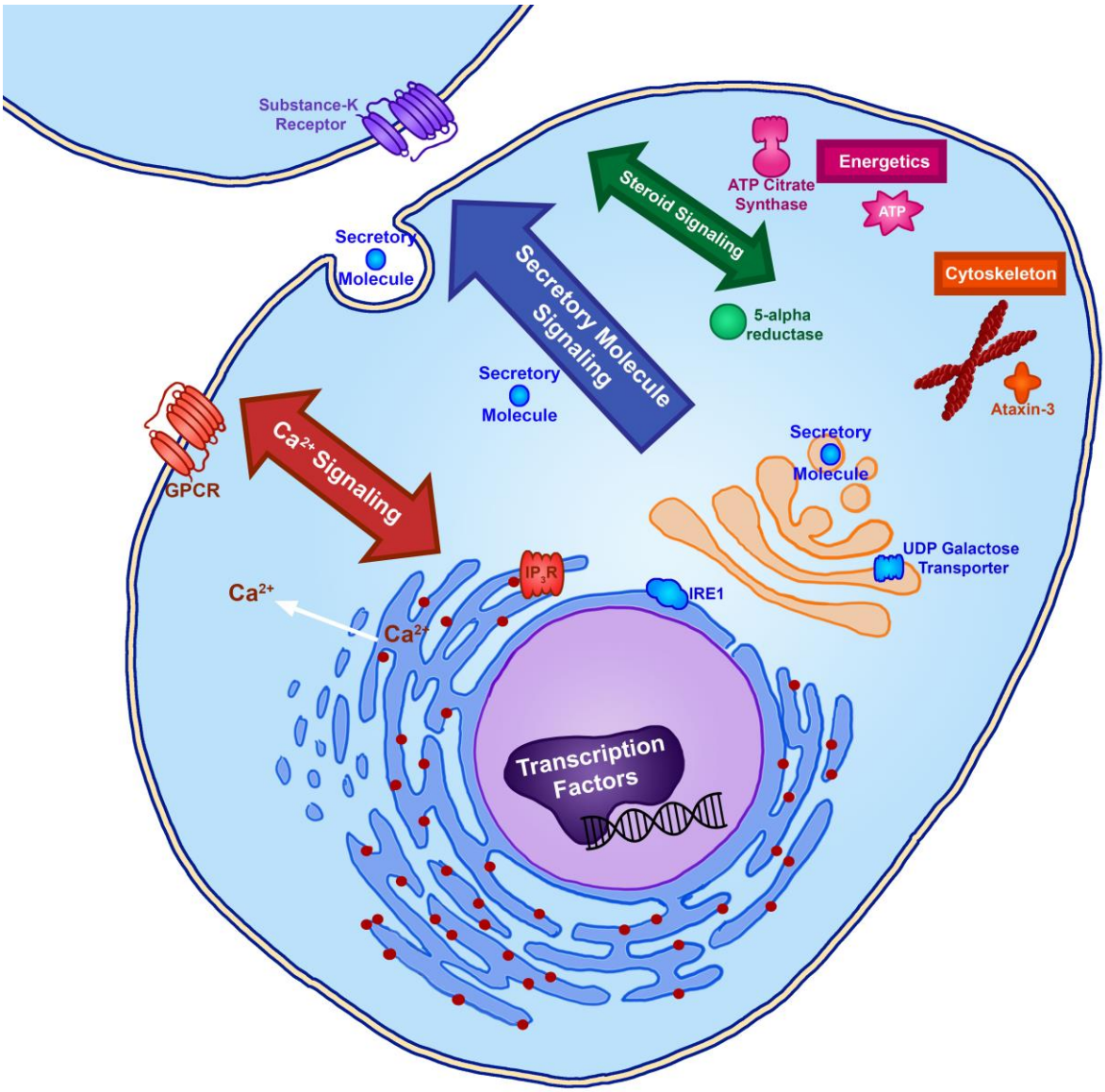


Figure 4-1. Summary figure indicating that our data shows ctenophores use multiple signaling pathways during regeneration. There is robust evidence for the use of the secretory molecule pathway, including peptide processing genes and predicted secretory molecules. Also indicated is the use of calcium signaling in the regenerating cell, which is an evolutionarily conserved mechanism in excited cells. Evidence was found for steroid signaling in the regenerating tissue. Energetic genes such as ATP producing proteins were shown to be up regulated as well as cytoskeleton and adhesion molecules. The data indicates transcription factors were also involved in the regeneration process. Our data show that these evolutionarily conserved signaling pathways and ctenophore specific predicted signaling molecules work in parallel to achieve ctenophore regeneration in this tissue.

Limb Regeneration	Vertebrate Regeneration	Neuronal Regeneration		Zebrafish Fin Regeneration	Muscle Regeneration	Wound Regeneration
BMP	BMP	AC	KLP-7	ACVR1	TFG-β	EGF
CTNNB1	MMP	AKT1	LIF	BMP	HGF	EGFR
FGF	Msx-1	APC	MAP3K1	CTNNB1	TNFα	FGFR
Fibrin	NOTCH1	ATF3	MAPK	FGF	IL-6	FGF
FN	NRG1	c-Jun	MEK-1	Hh	NOS	HGF
MMP	Substance P	CCPP-6	MKK-4	IGF		HGFR
Msx-1	TF	CEBP-1	MLK-1	mTOR		IGF
Msx-2		CED-3	mTOR	NOTCH1		IL-1
Msx-9		CED-4	NCID	RA		IL-1R
TFG-β		CLASPs	NOTCH1	Wnt		IL-6
Wnt		CNTF	PI3K			KGF
		DLK-1	PTEN			KGFR
		EFA-6	Smad			MCP-1
		ERK	SOC3			PDGF
		GSK3	SOX11			PDGFR
		HDAC5	STAT3			SDF1
		IL-6	SVH-1			SDF1R
		Jak	SVH-2			SOC3
		JIP3	Tuba1a			TGFβ
		JNK				TGFβR
		KLF				TNFα
						VEGF
45%	57%	20%		30%	20%	36%
9%	0%	28%		20%	40%	41%
27%	29%	43%		10%	40%	23%
18%	14%	10%		40%	0%	0%

Key
Not Present in *Mnemiopsis*
Present in *Mnemiopsis*, Not Significantly Changed in Regeneration
Present in *Mnemiopsis*, Up-regulated in Regeneration
Present in *Mnemiopsis*, Down Regulated in Regeneration

Figure 4-2. Summary figure showing relationship between genes previously found to be involved in regeneration and their expression in ctenophore regeneration. We studied genes complied from limb regeneration by Yokoyama 2008 (6), cellular plasticity in vertebrate regeneration by Odelberg 2005 (7), neuronal regeneration by Saijilafu et al 2013 (10) and Lu et al 2014 (9), zebrafish fin regeneration by Wehner and Weidinger 2015 (5), muscle regeneration by Charge et al 2004 (52), and wound healing by Eming et al 2014 (8). These genes are represented here in this table. The colors represent the result of our analysis. If the gene is grey, it is not present in the *Mnemiopsis* gene models. If the gene is black, it is present in the gene models but not differentially expressed in regeneration. If the gene is green it is up regulated in regeneration. If the gene is blue it is down regulated in regeneration. The percentages under each column represent the percentages for each group of genes in that category. For example, in limb regeneration 45% of the genes in the limb regeneration catagory are not present in the *Mnemiopsis* gene models. To see more information about these genes see *Supplementary Data*.

CHAPTER 5
FUTURE DIRECTIONS

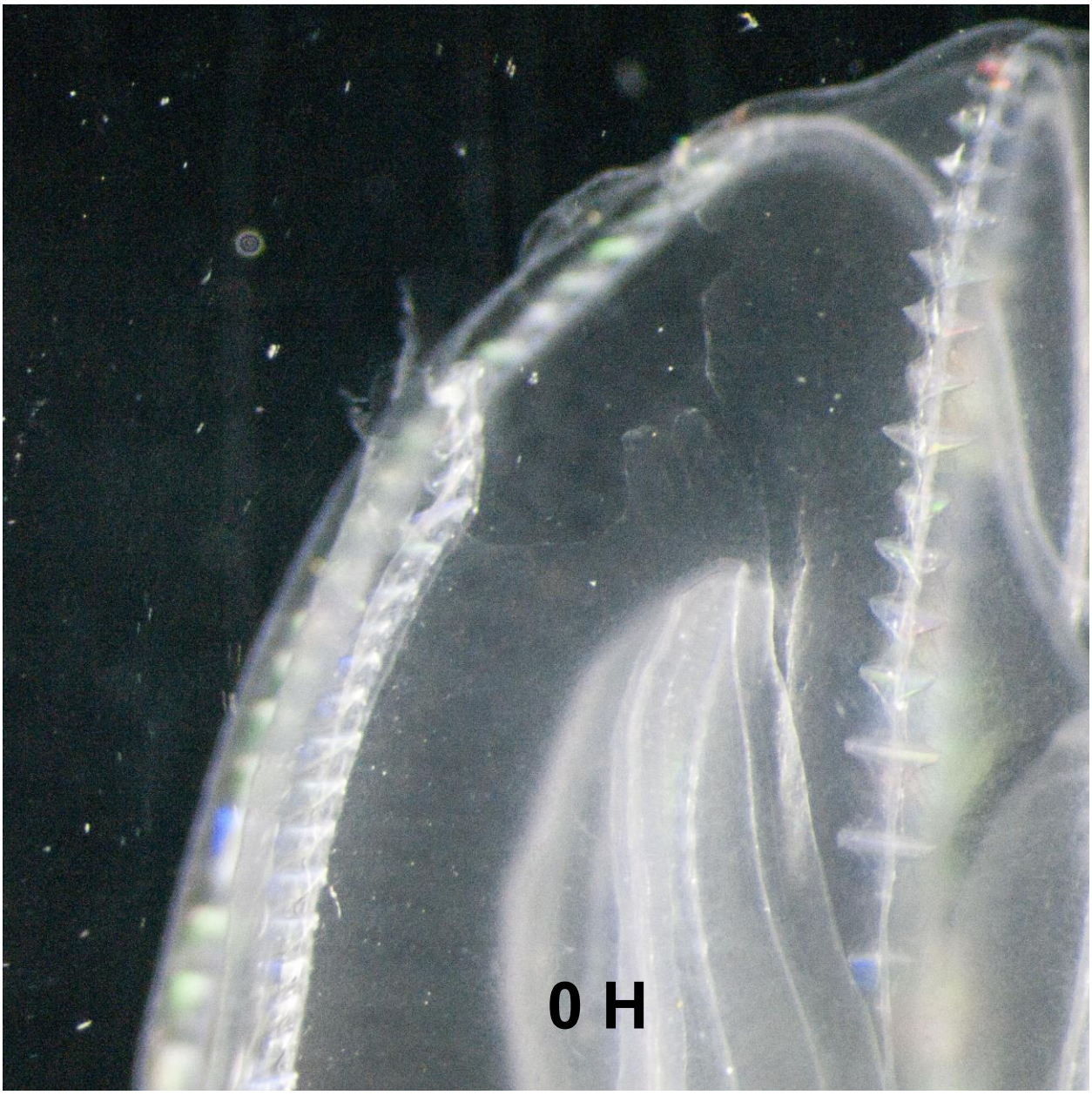
There are many avenues of future work from this study. The aboral organ removal experiments will be performed with replicates and control tissue and sequenced using Illumina sequencing to produce a larger and more robust data set. This will give us much insight into neural regeneration in this animal and how it differs from other regeneration types. Detailed immunoflurence visualization of regeneration will also be performed. As studies started by Andrilenas and Moroz in 2010, FMRFamide and α-tubulin will be used to characterize epithelial neural net cells during different phases of regeneration. Also in different phases of regeneration EdU (5-ethynyl-2'-deoxyuridine) labeling will be used during regeneration to determine DNA-replicating cells, as used in by Alie et al (42).

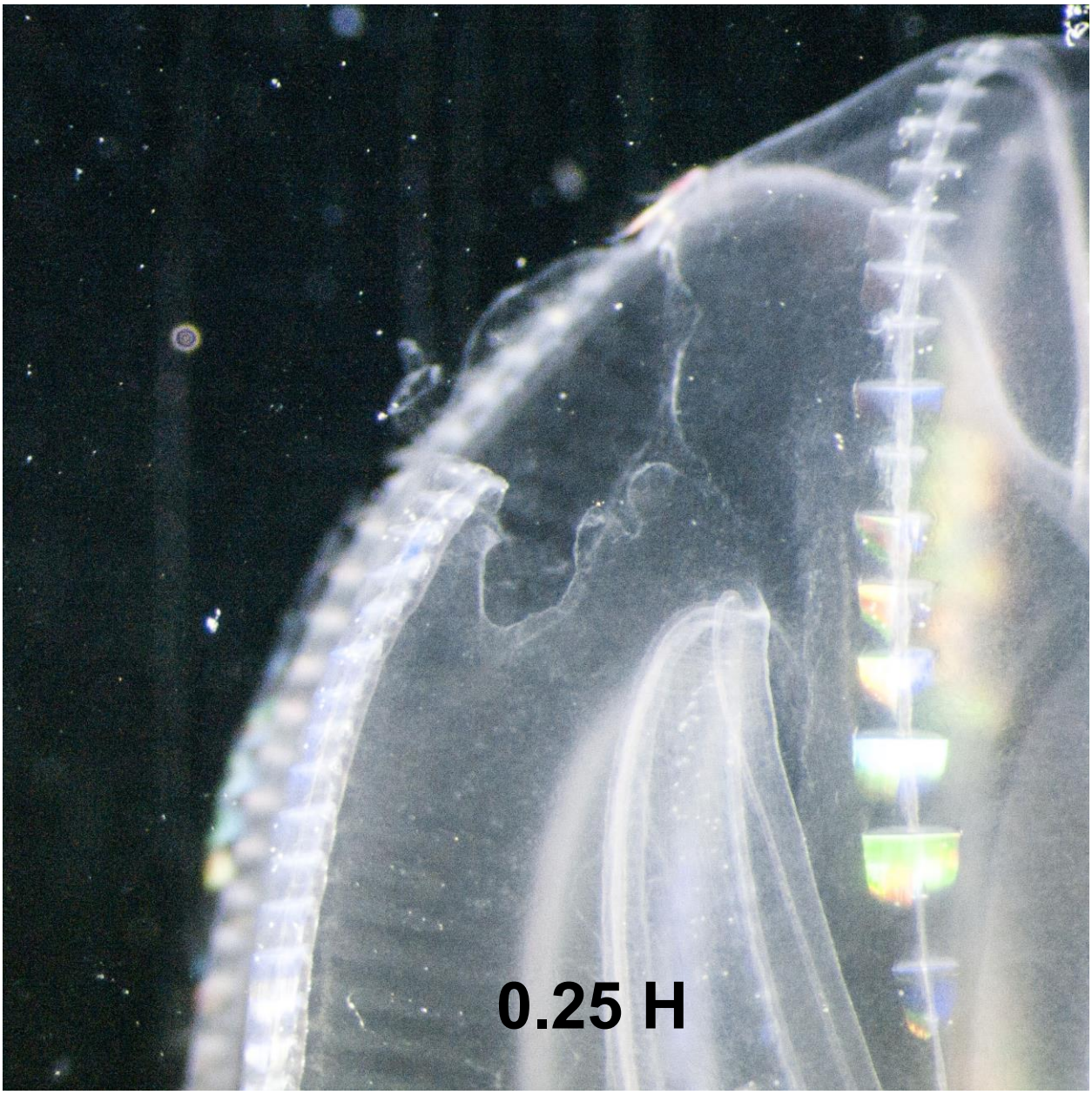
Functional studies will also be performed. The results indicate that secretory peptides are a key player in regeneration signaling. We will have 10 of these secretory peptides synthesized and in turn we will perform regeneration experiments with them. We will compare regeneration time of normal ctenophores, to those that we have incubating in a bath of the signaling molecule. Any that show promise might become candidates for new therapy drugs trials.

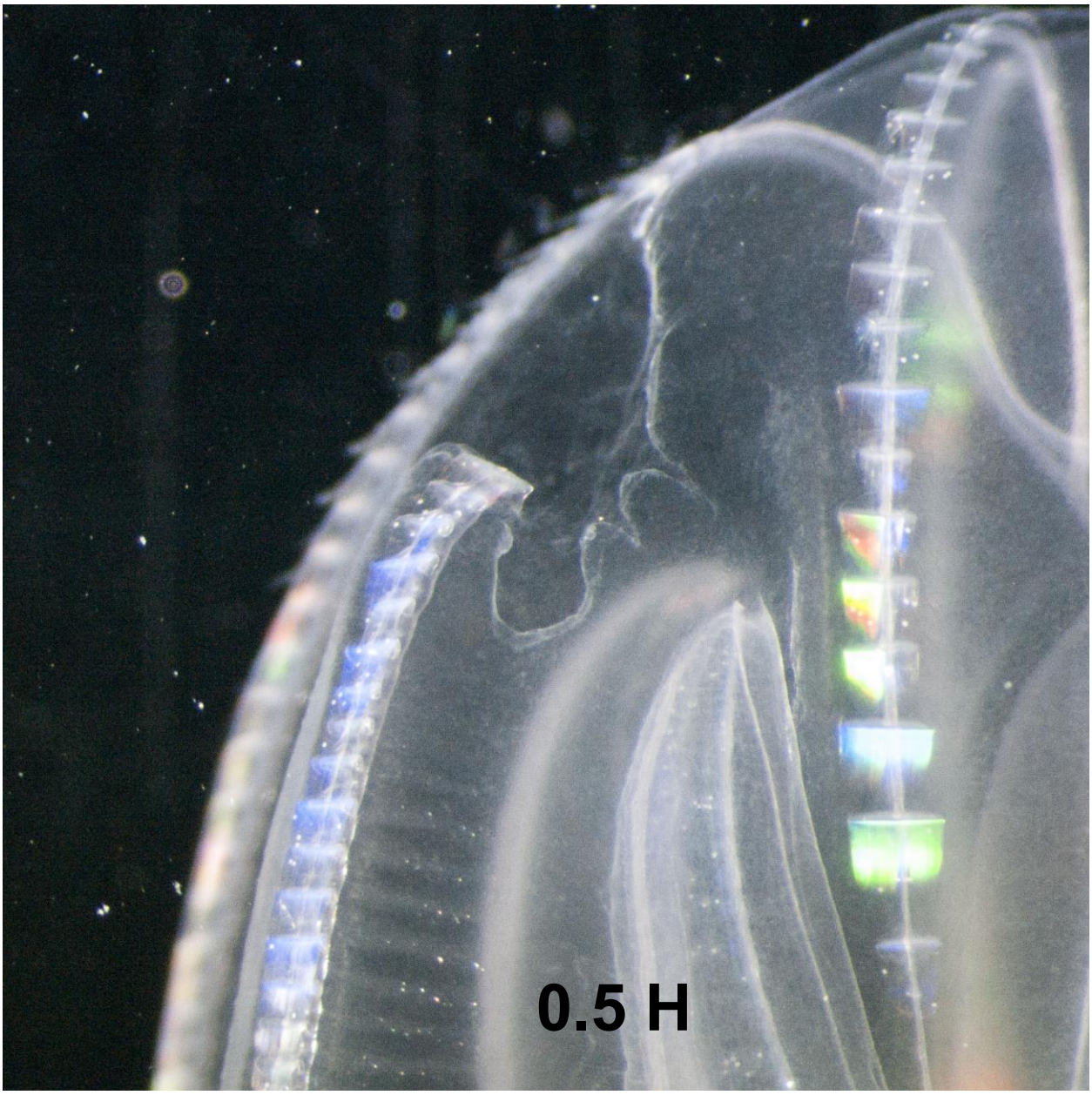
Another important future direction is to analyze the methylome during regeneration. While the genes play a large role in regeneration, the methylation pattern of these genes coordinates gene expression, as has been shown in other animals (71-73). We have started preliminary work on this and have sequenced the methylome of a control animal and a regenerating ctenophore. See Appendix Figure A-4 and

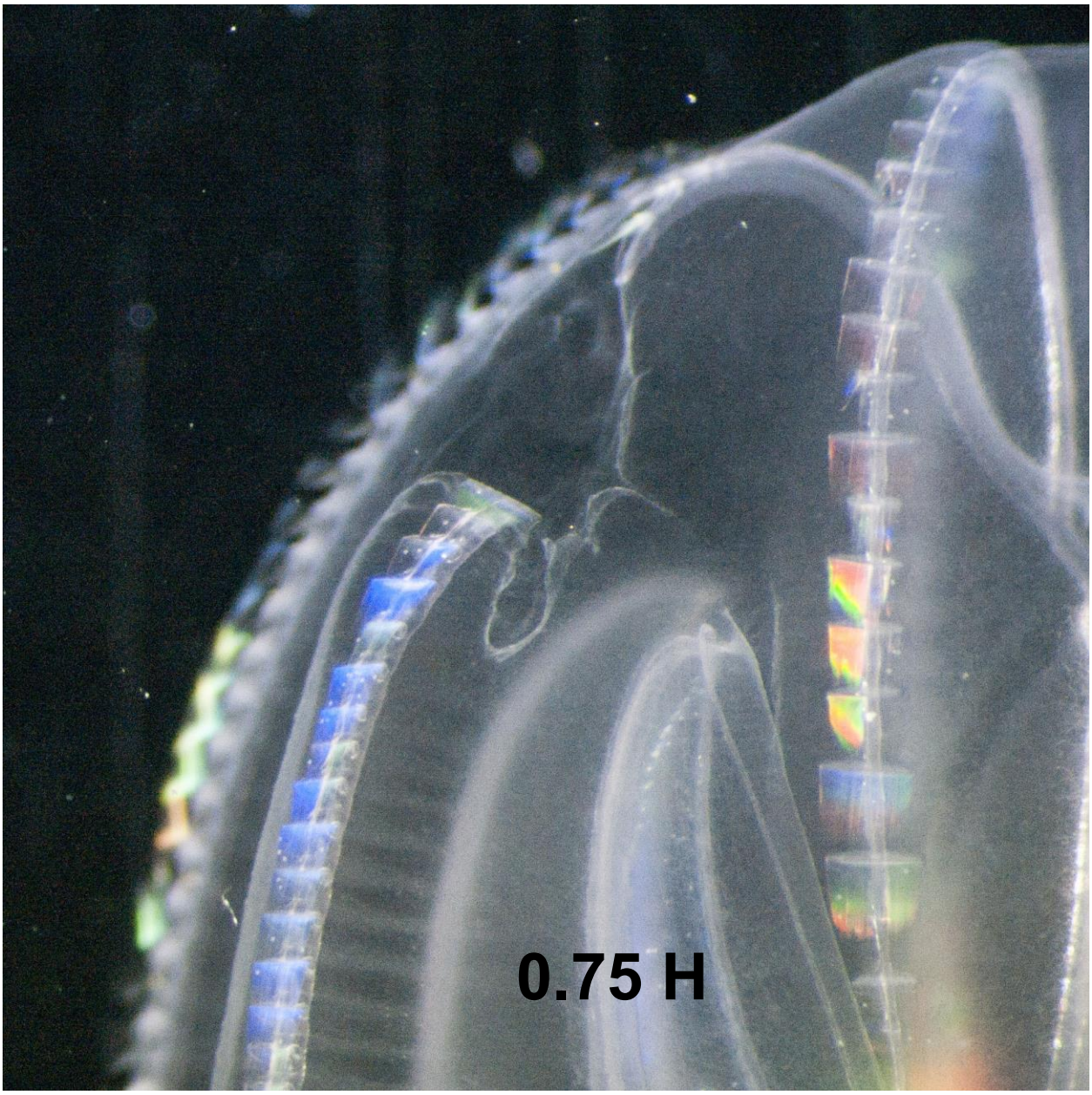
Supplementary Data for our current analysis. Future work would include more replicates of this study.

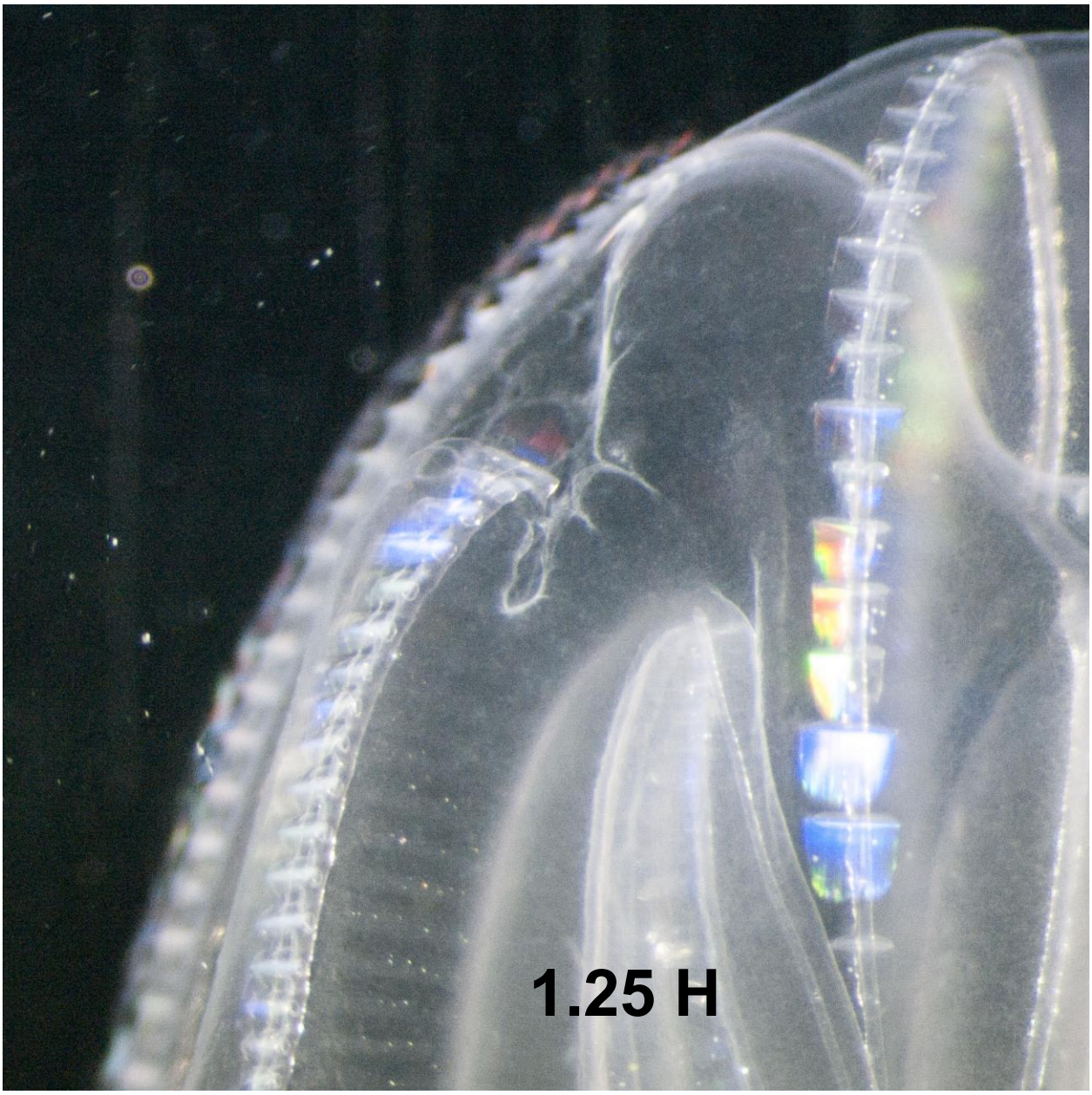
APPENDIX
SUPPLEMENTAL FIGURES

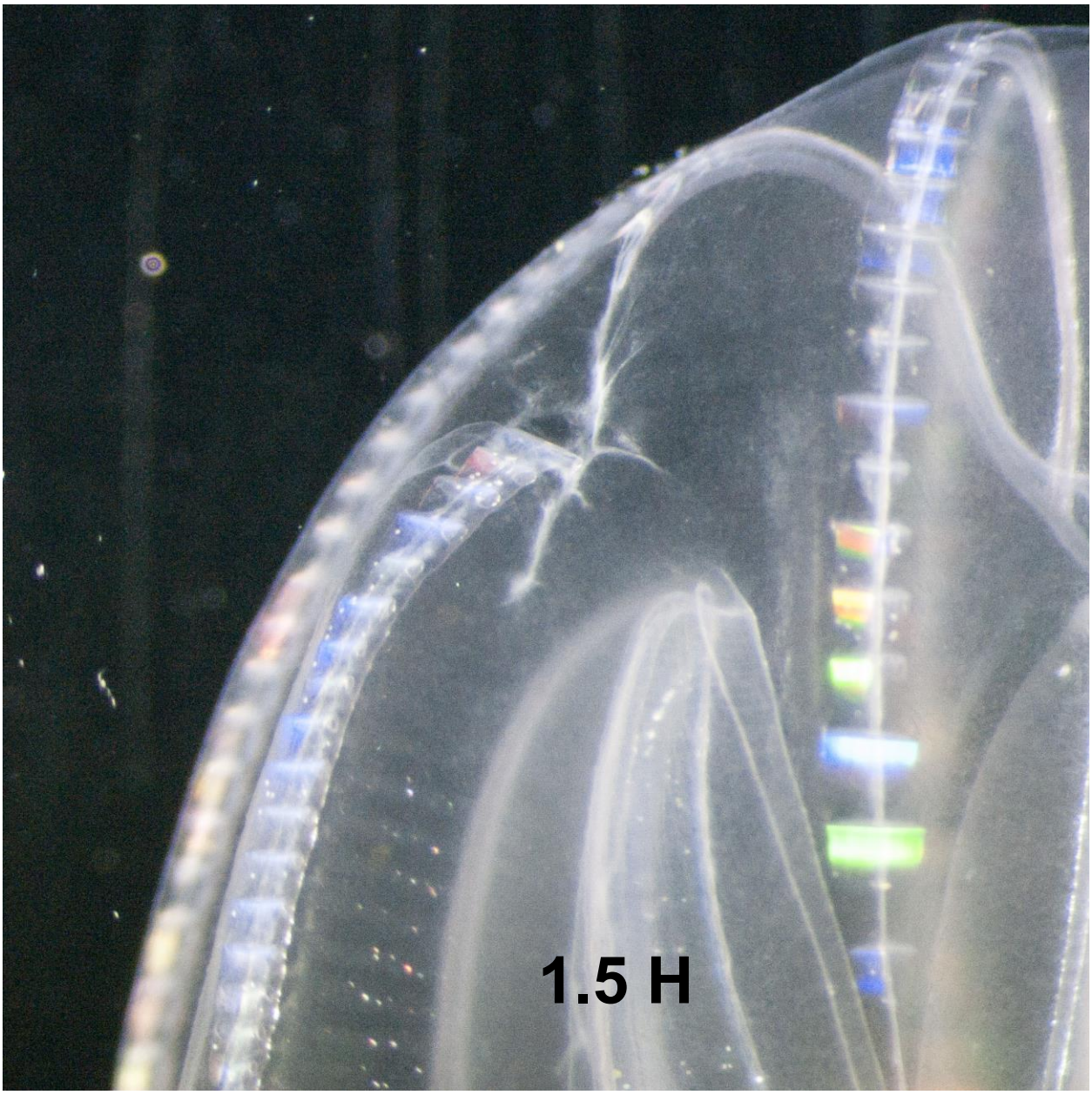

0 H

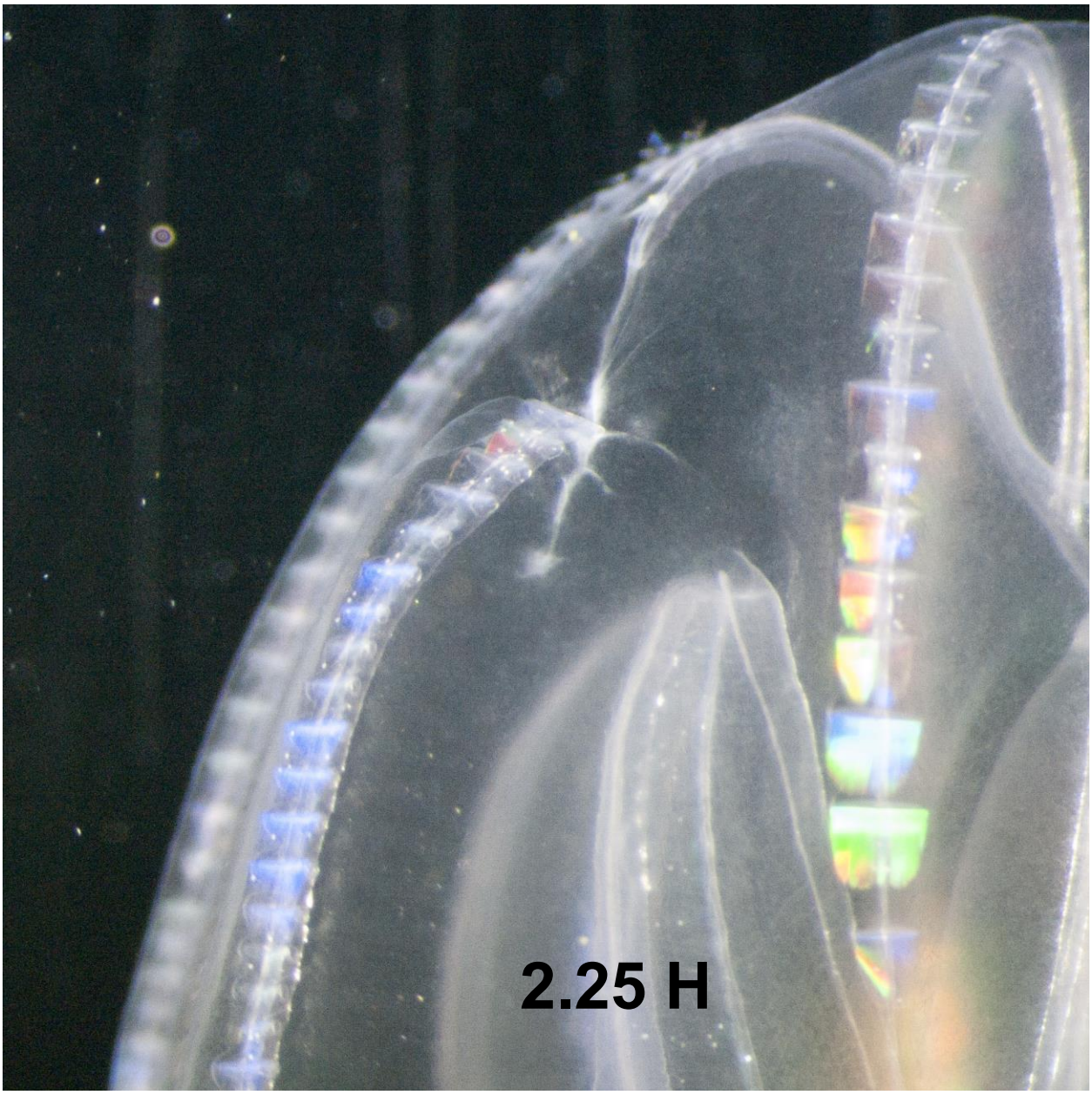

0.25 H

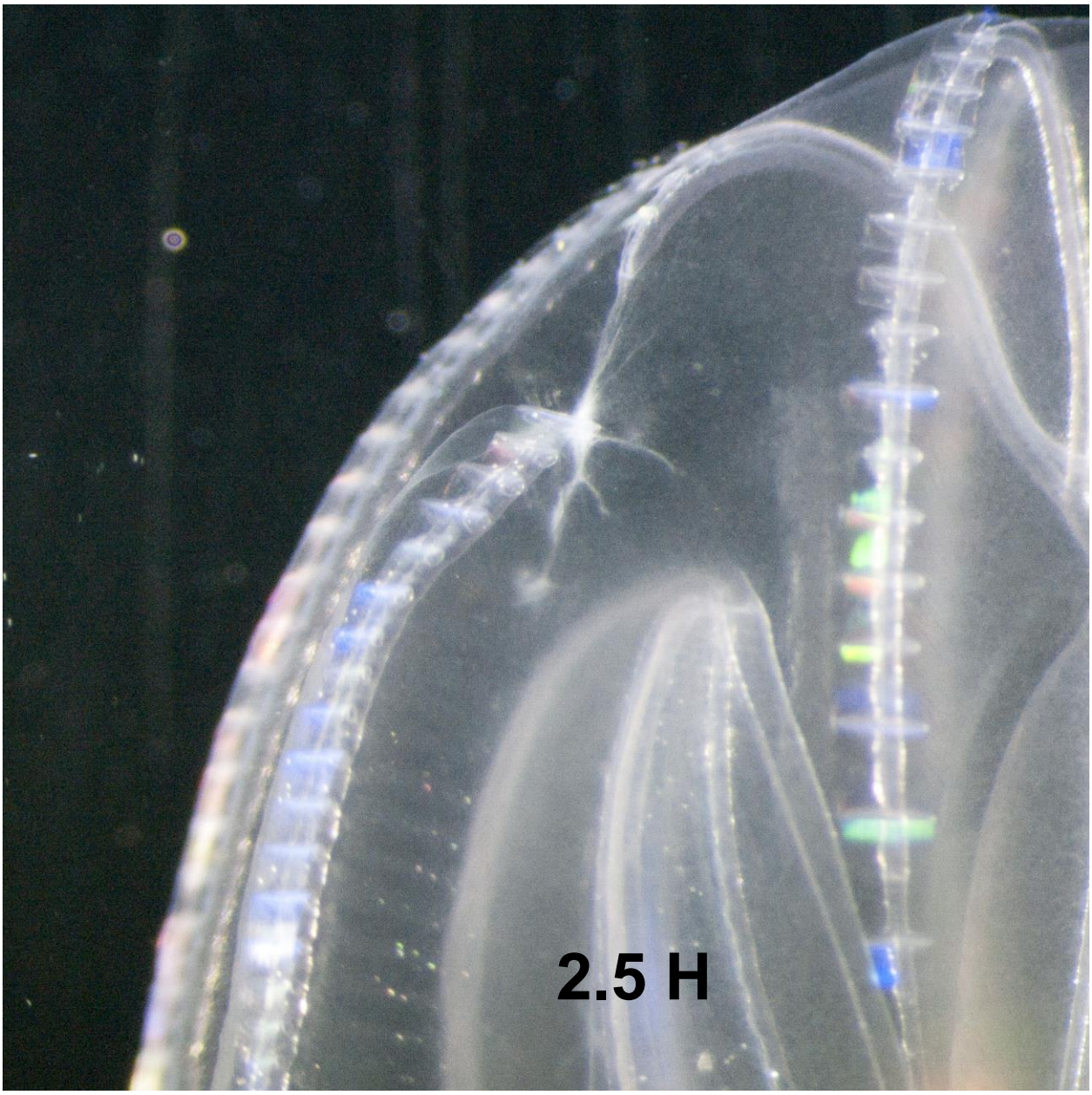

0.5 H

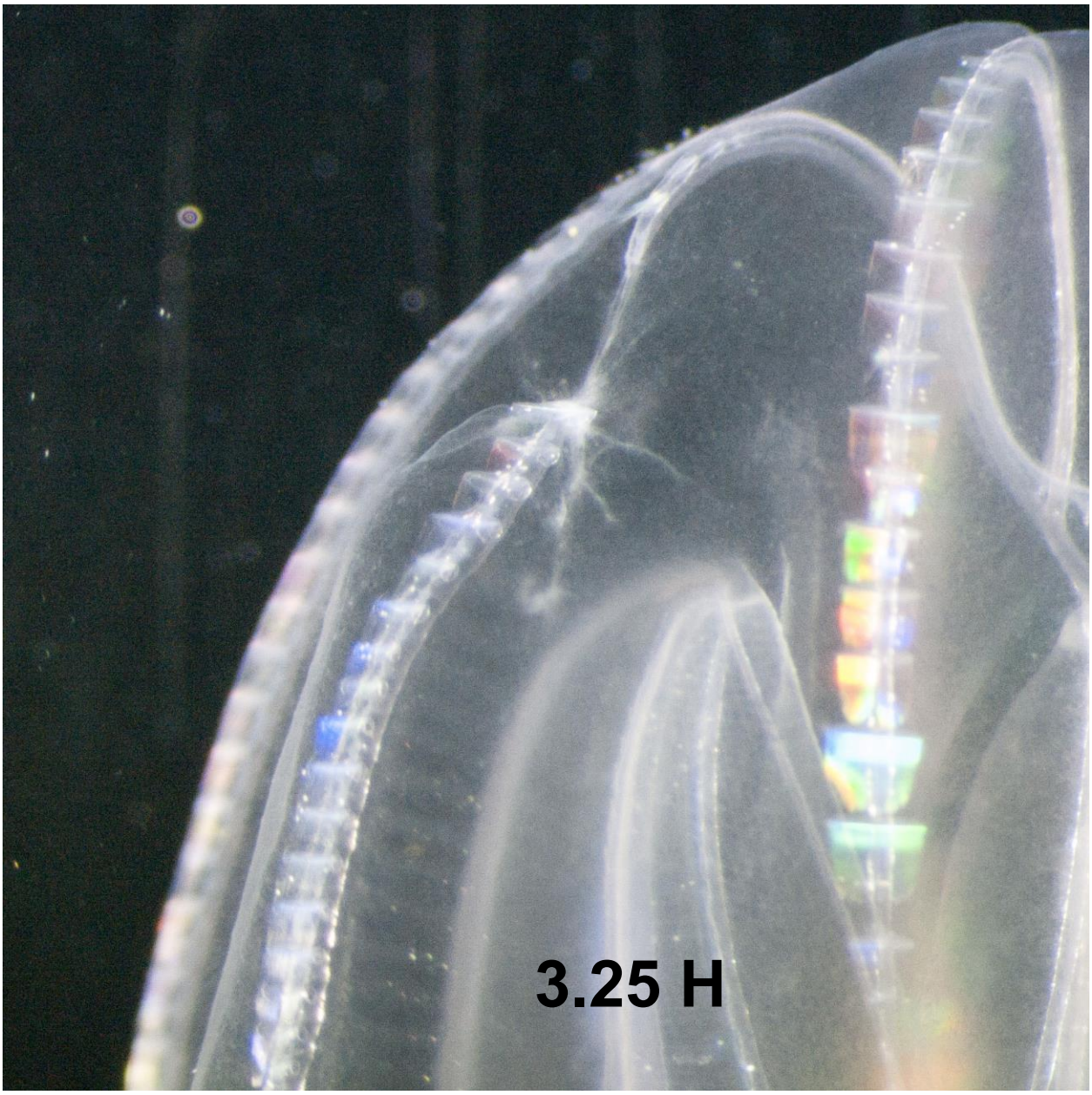

0.75 H

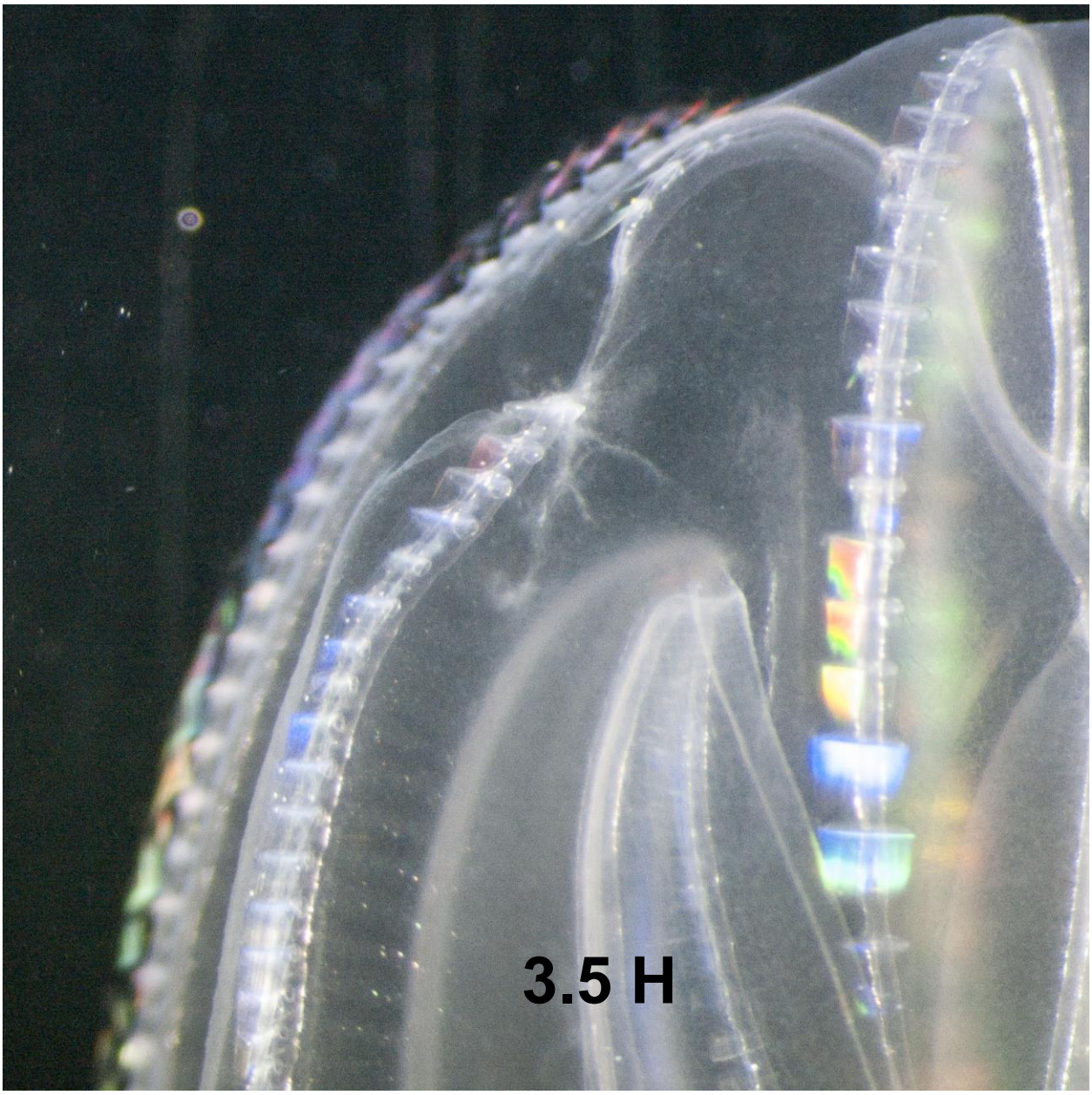

1.25 H

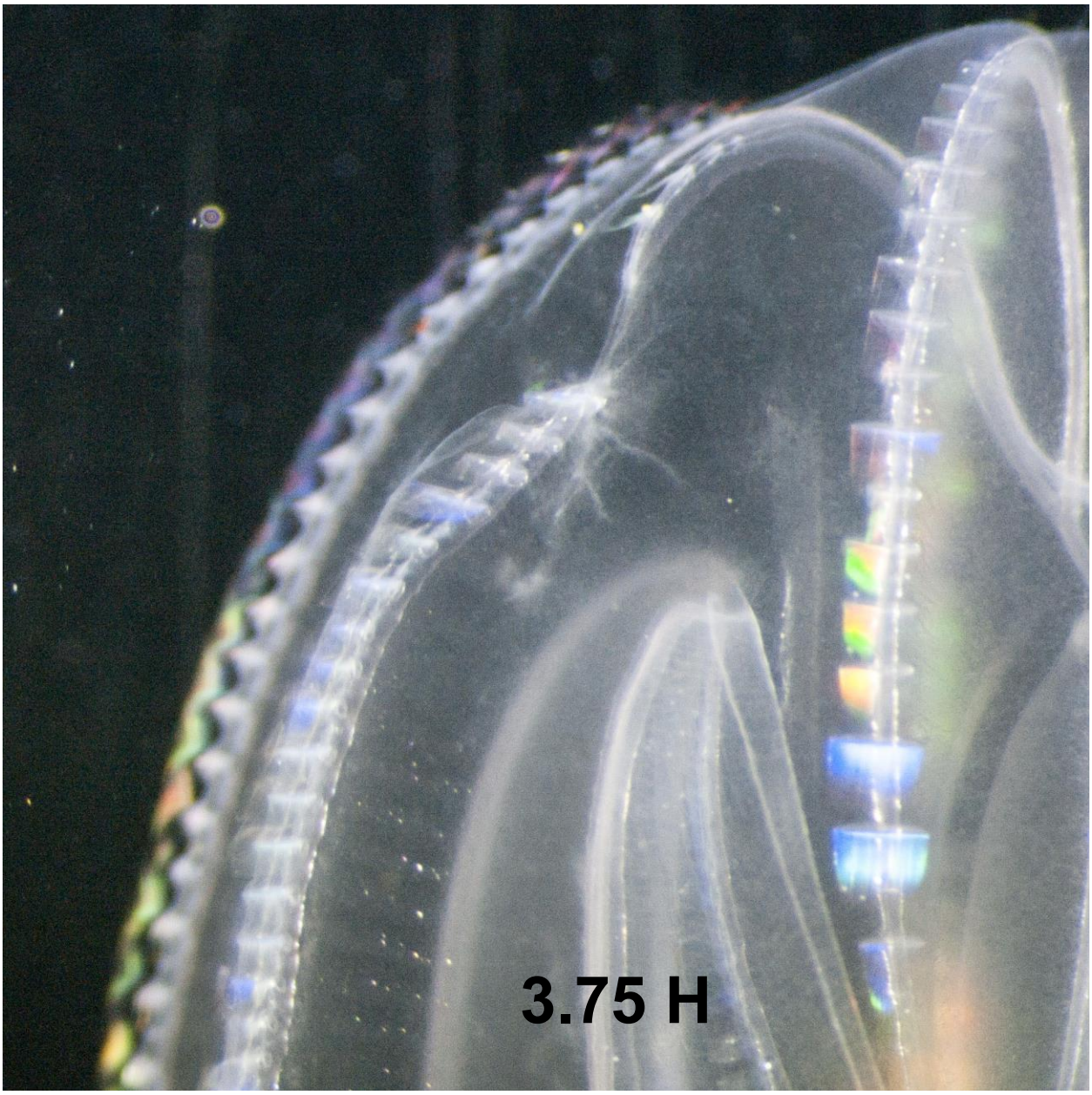

1.5 H

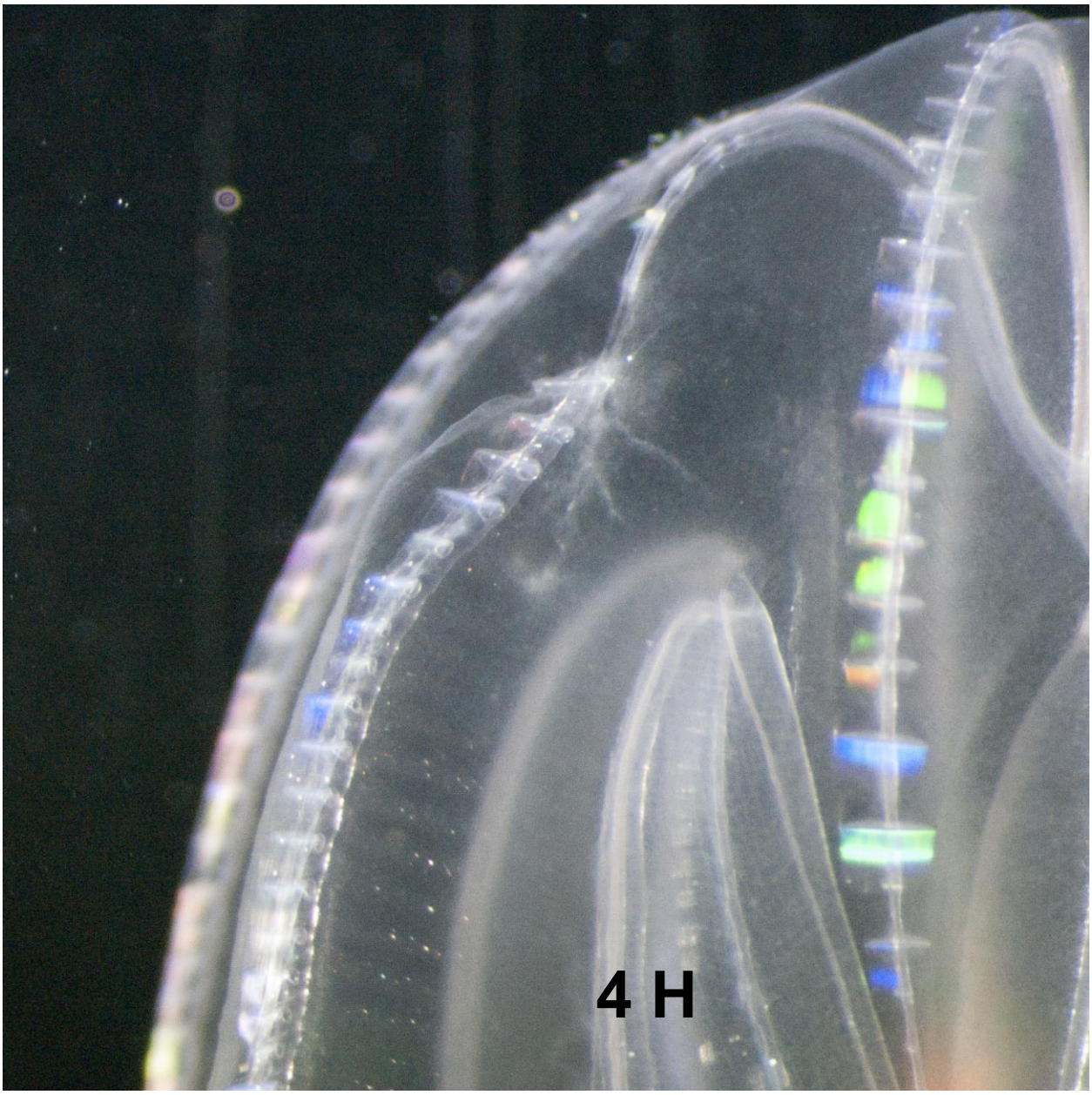

2.25 H

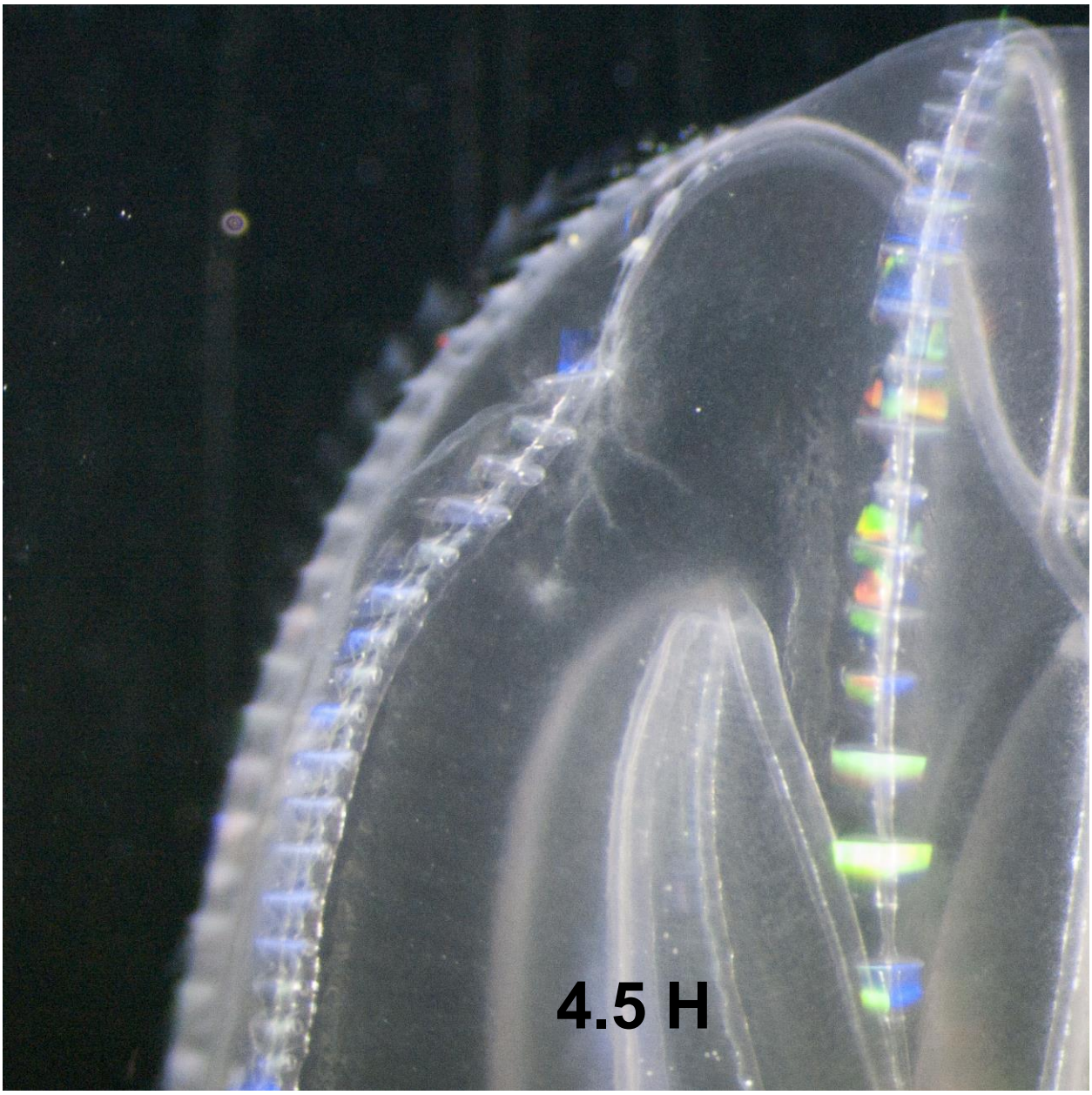

2.5 H

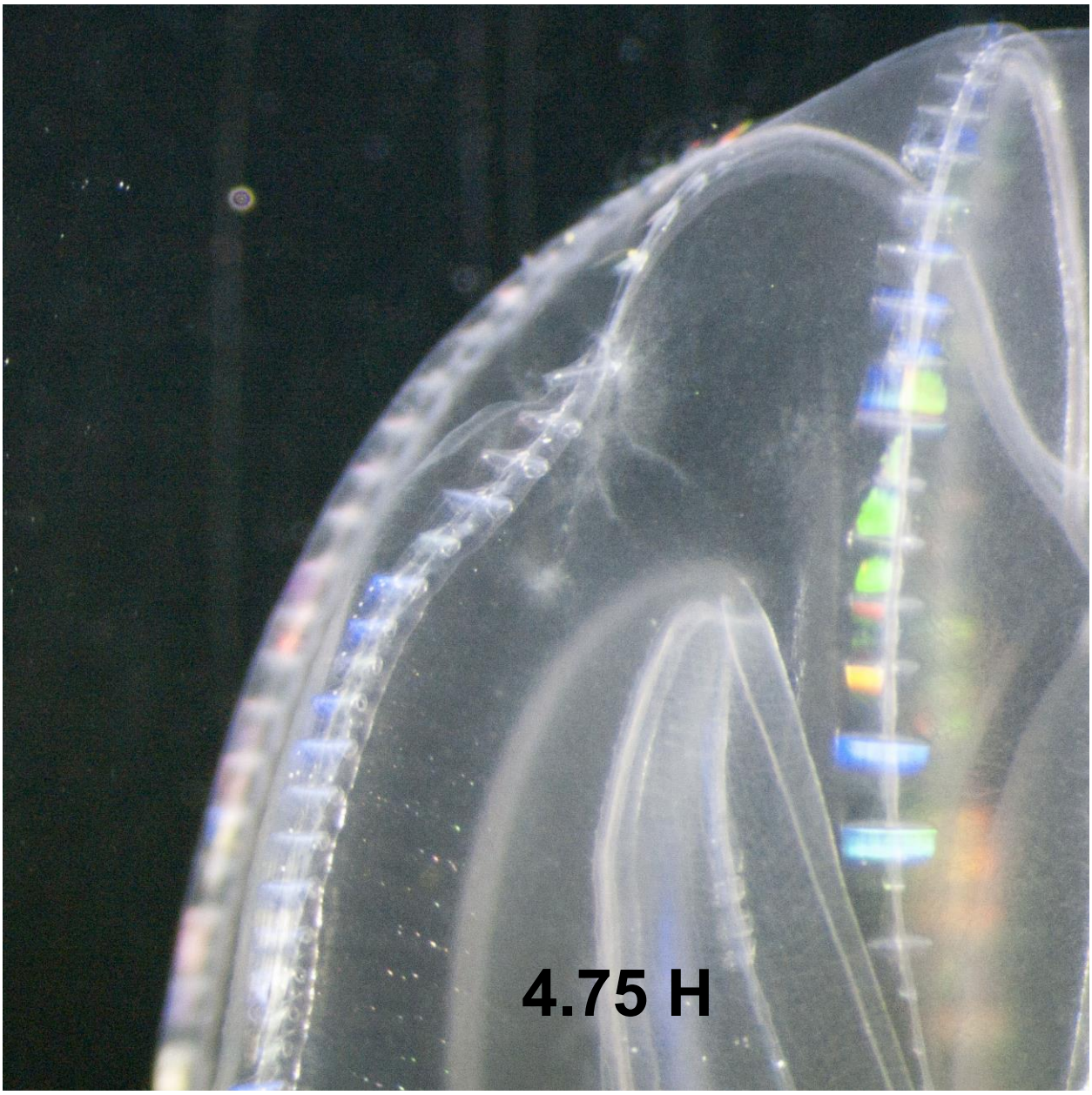

3.25 H

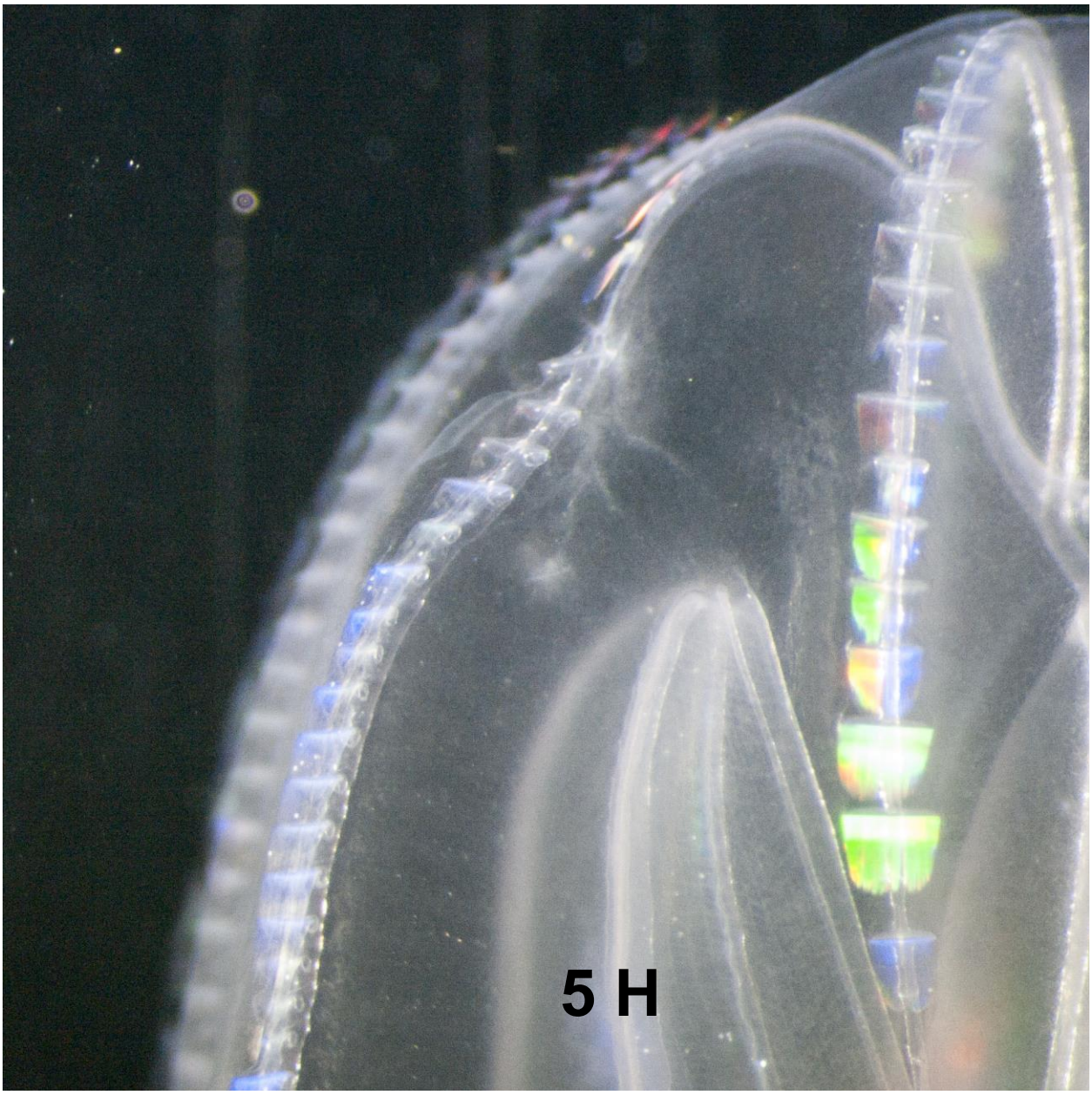

3.5 H

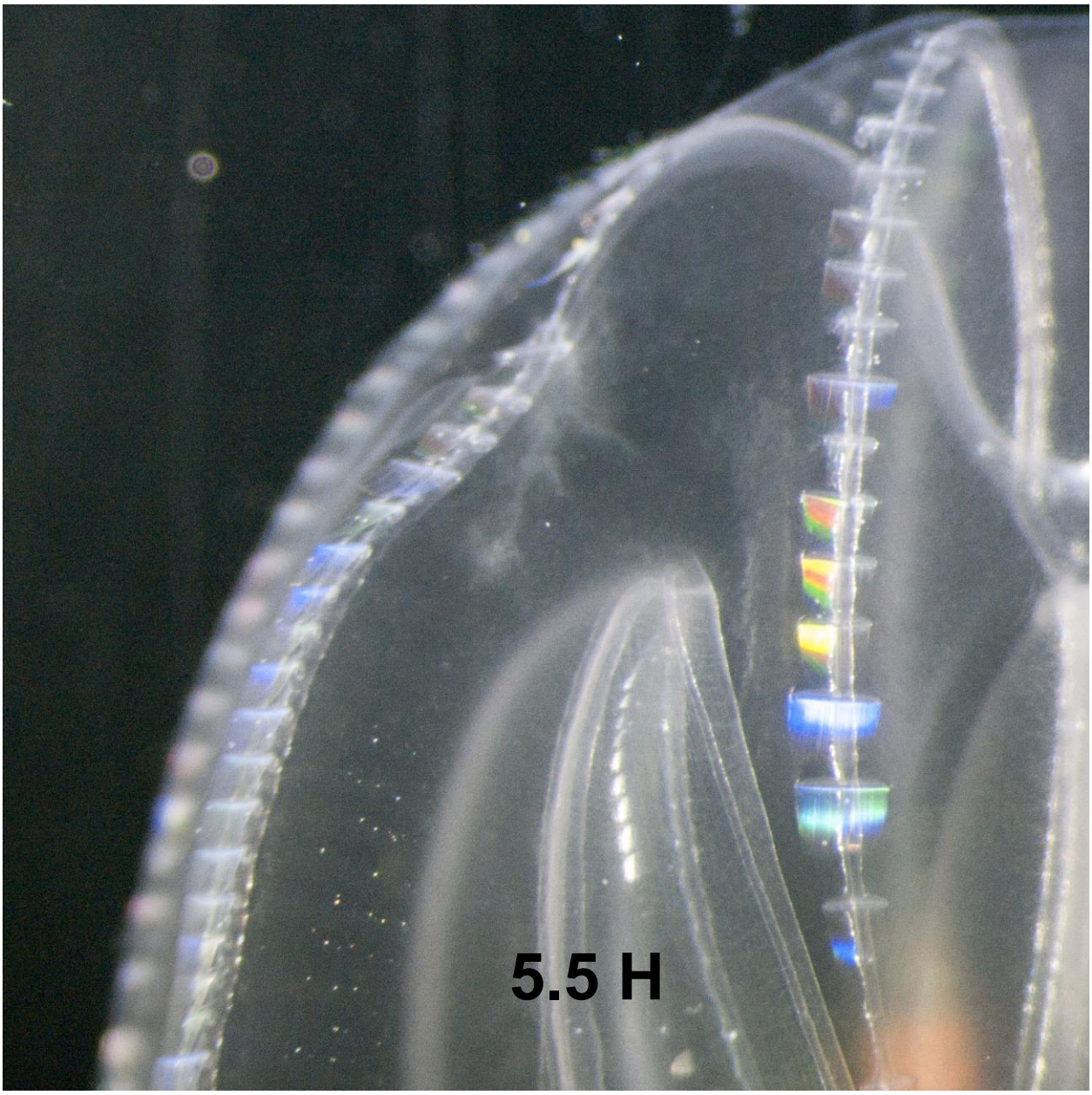

3.75 H

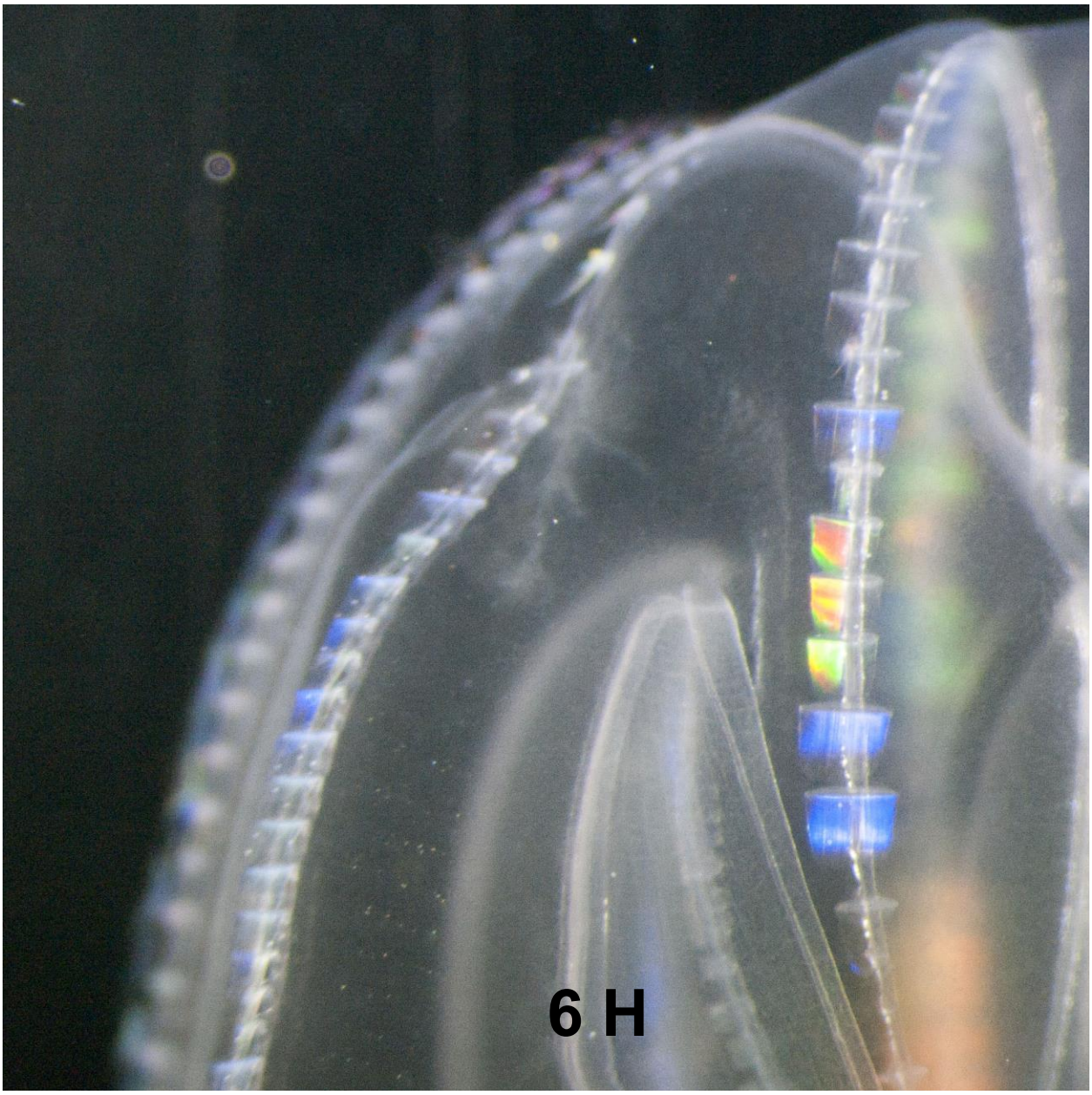

4 H

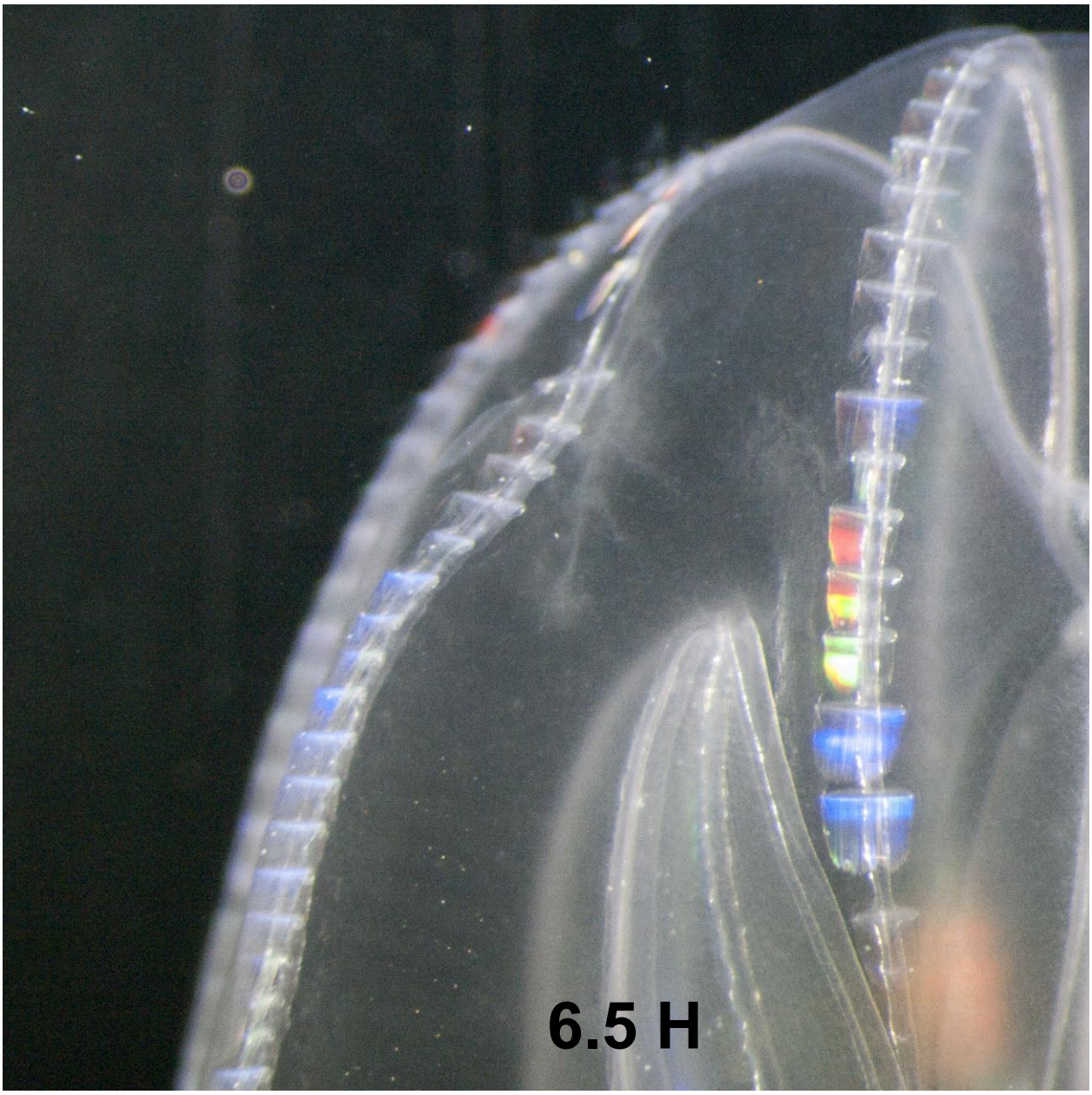

4.5 H


4.75 H


5 H


5.5 H


6 H


6.5 H

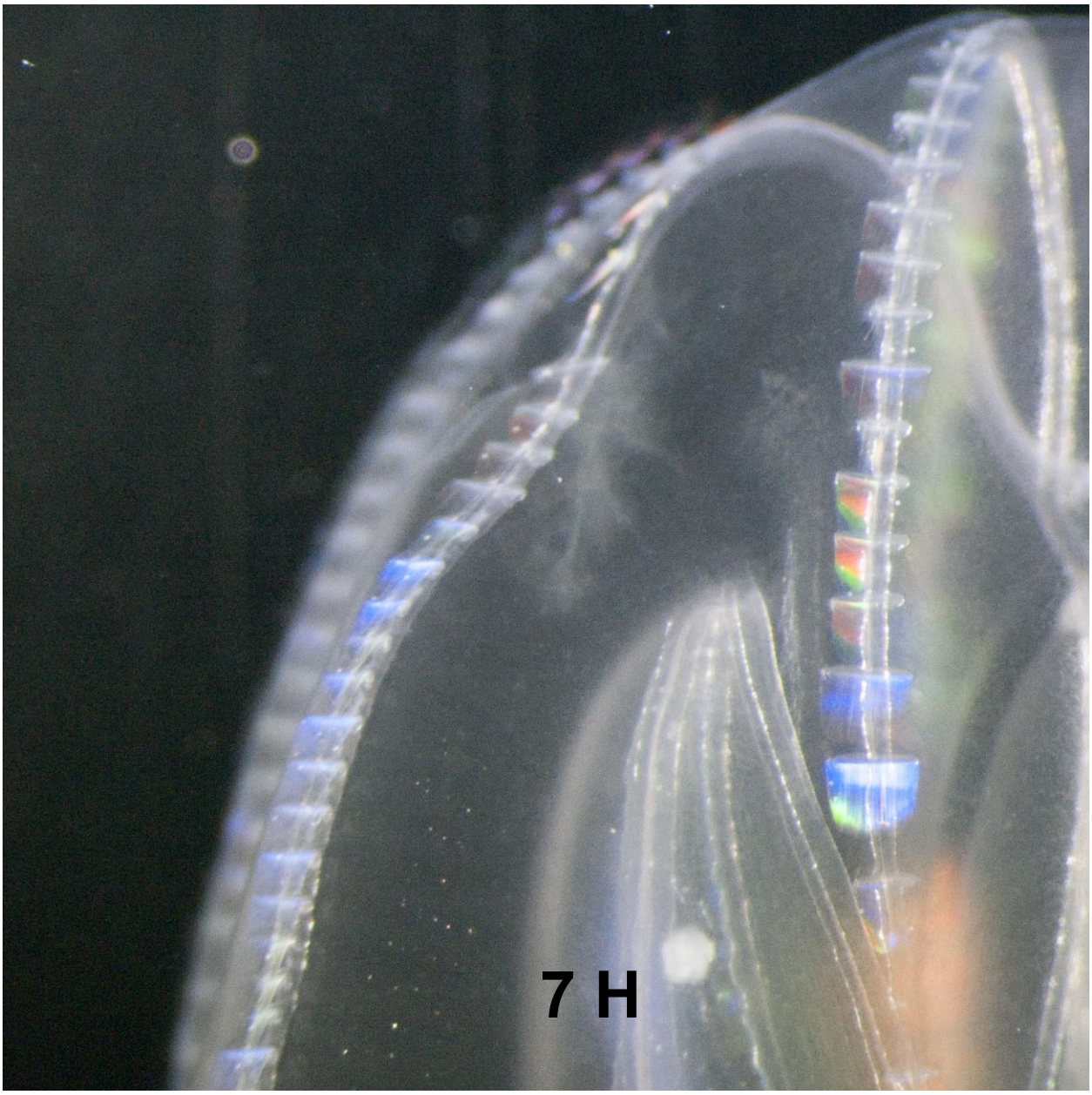


Figure A-1. Large version of ctenophore regeneration pictures. These 20 pictures were taken over seven hours. The ctenophore was pinned to a dish of silgard to reduce movement. The pictures were taken using a Pentax K-7 DSLR mounted to a dissecting scope using a microscope adapter. The ctenophore was injured using forceps and imaged approximately every 15 to 30 minutes. Fresh seawater was carefully changed in the dish using a pipette. Sea water was room temperature.

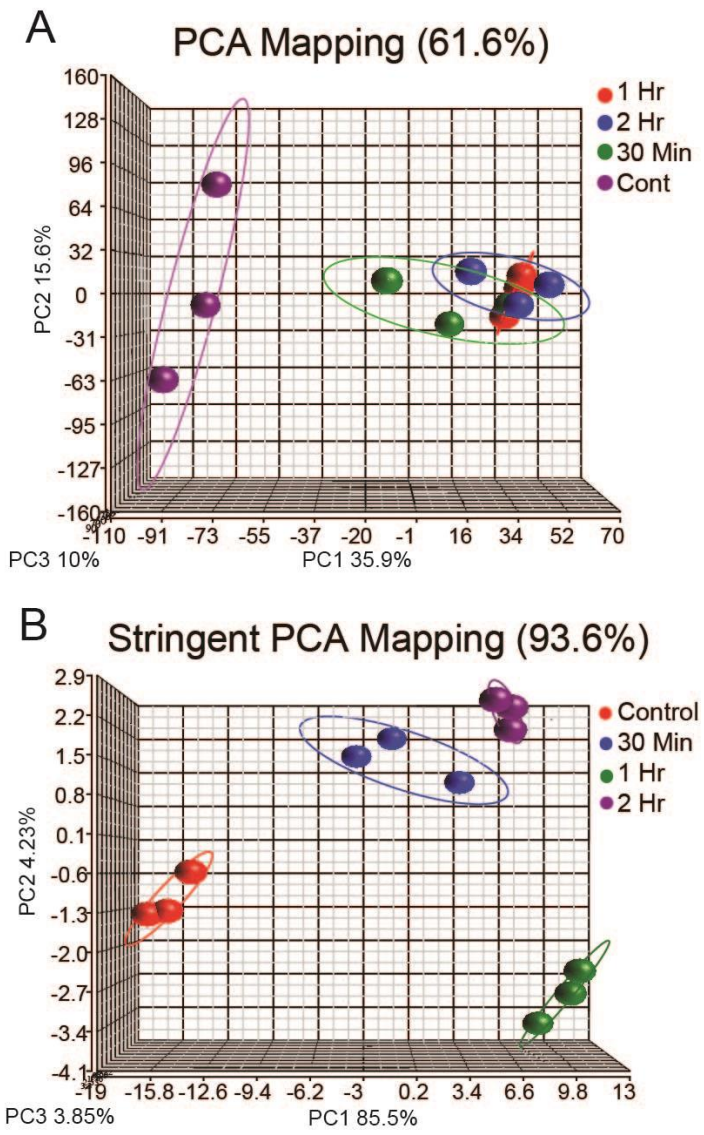


Figure A-2. The principal component analysis (PCA) mapping for the Illumina data. A) The Illumina RNA-Seq data clustered relatively well, showing 61.6% overall mapping and each treatment clustered in small groups, though some overlapped. B) A more stringent cutoff of the Illumina data was used for the following gene list described in the section, *Multiple Signaling Pathways Initiate Regeneration in Ctenophores*, as defined as FDR < 0.005, fold change > 5 or < -5. The PCA for these groups of genes are more tightly clustered, showing 93.3% mapping and distinct groups.

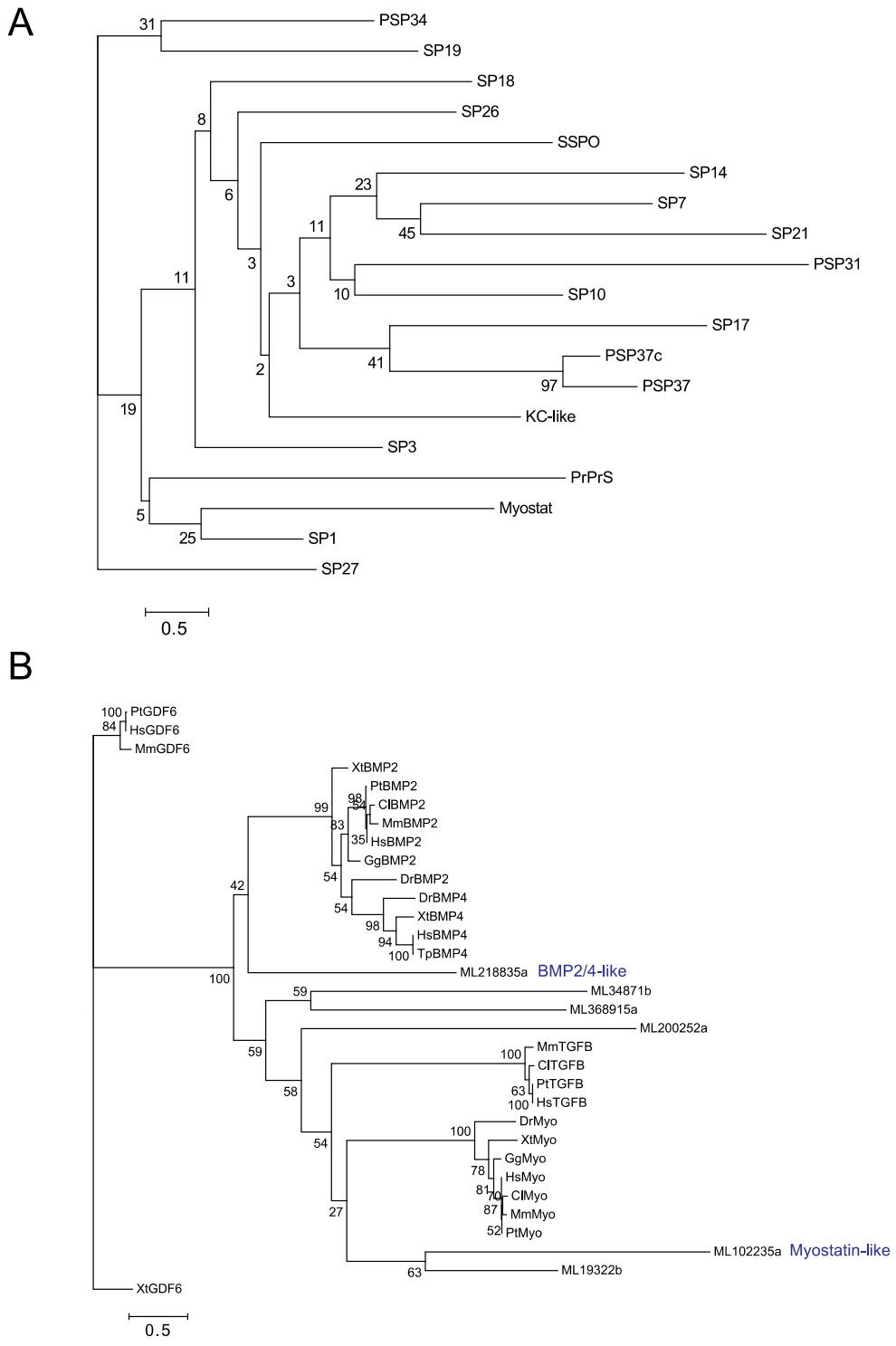


Figure A-3. Phylogenic trees made by RaxML (74) A) Predicted secretory peptides show very little homology between each other B) TGF-B family peptides plus ML sequences, two ML sequences termed “BMP2/4-like” and “Myostatin-like” in blue. Alignment was done using MUSCLE (75). No g-block trimming was performed.

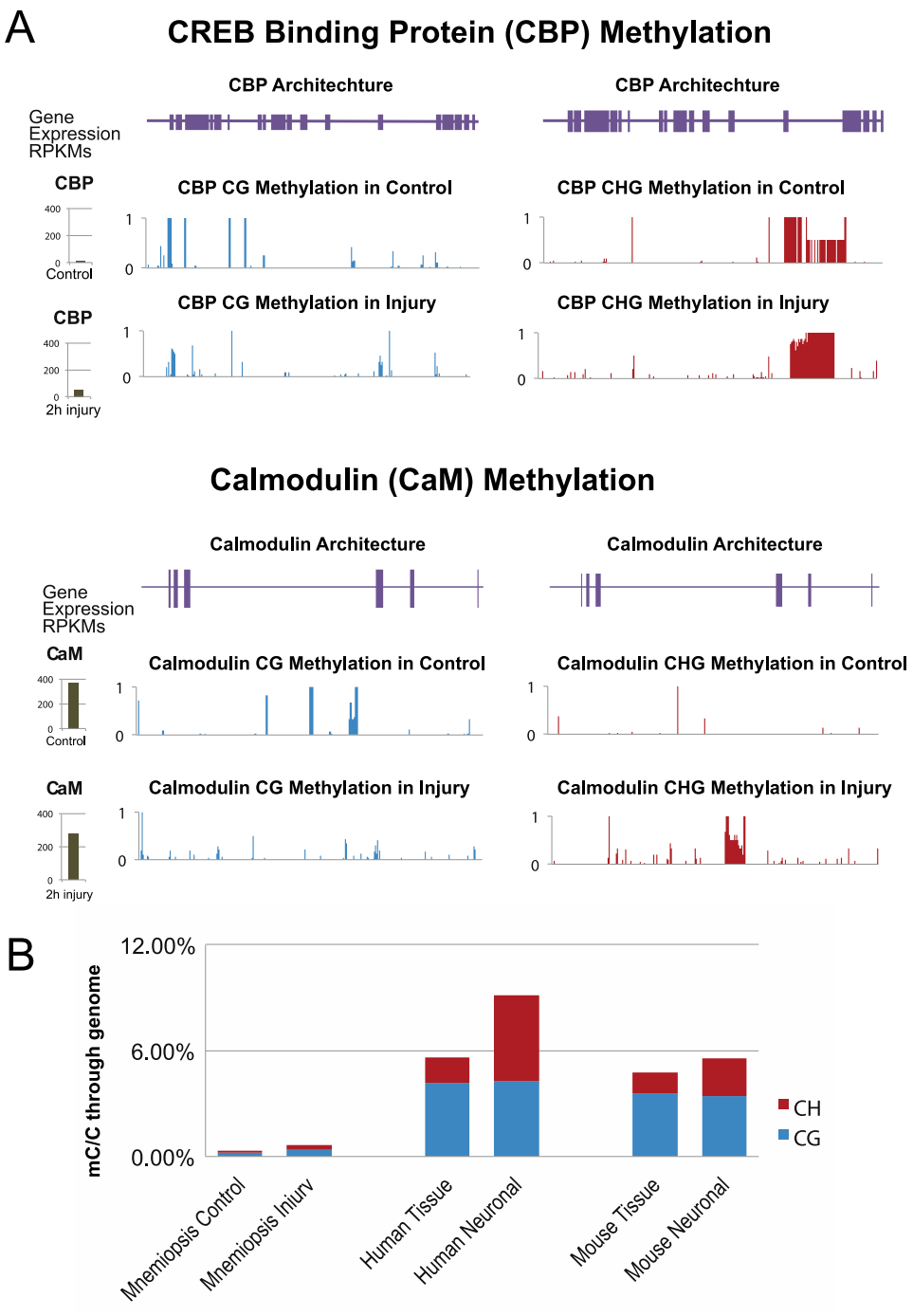


Figure A-4. Methylation analysis A) Individual gene analysis of CBP and CaM methylation at CG sites and CHG sites. Intron / exon organization shown in dark purple. Blue and red graphs show individual methylation sites along the gene. B) Overall C methylation in ctenophores compared to humans and mouse methylation. Less methylation in general in ctenophores compared to mammals. Methylation data produced using MOABs (76) and dibig_tools and help from Dr. Alberto Riva.

LIST OF REFERENCES

1. Alvarado AS & Tsonis PA (2006) Bridging the regeneration gap: genetic insights from diverse animal models. *Nature reviews. Genetics* 7(11):873-884.

2. Bely AE (2010) Evolutionary loss of animal regeneration: pattern and process. *Integr Comp Biol* 50(4):515-527.

3. Bely AE & Nyberg KG (2010) Evolution of animal regeneration: re-emergence of a field. *Trends Ecol Evol* 25(3):161-170.

4. Morgan TH (1901) *Regeneration, by Thomas Hunt Morgan* (The Macmillan Company;, New York).

5. Wehner D & Weidinger G (2015) Signaling networks organizing regenerative growth of the zebrafish fin. *Trends Genet* 31(6):336-343.

6. Yokoyama H (2008) Initiation of limb regeneration: the critical steps for regenerative capacity. *Dev Growth Differ* 50(1):13-22.

7. Odelberg SJ (2005) Cellular plasticity in vertebrate regeneration. *Anat Rec B New Anat* 287(1):25-35.

8. Eming SA, Martin P, & Tomic-Canic M (2014) Wound repair and regeneration: mechanisms, signaling, and translation. *Sci Transl Med* 6(265):265sr266.

9. Lu Y, Belin S, & He Z (2014) Signaling regulations of neuronal regenerative ability. *Curr Opin Neurobiol* 27:135-142.

10. Saijilafu, Zhang BY, & Zhou FQ (2013) Signaling pathways that regulate axon regeneration. *Neurosci Bull* 29(4):411-420.

11. Moroz LL, Kocot KM, Citarella MR, Dosung S*, et al.* (2014) The ctenophore genome and the evolutionary origins of neural systems. *Nature* 510(7503):109-114.

12. Whelan NV, Kocot KM, Moroz LL, & Halanych KM (2015) Error, signal, and the placement of Ctenophora sister to all other animals. *Proceedings of the National Academy of Sciences* 112(18):5773-5778.

13. Ryan JF, Pang K, Schnitzler CE, Nguyen AD*, et al.* (2013) The genome of the ctenophore Mnemiopsis leidyi and its implications for cell type evolution. *Science* 342(6164):1242592.

14. Haddock SD (2004) A golden age of gelata: past and future research on planktonic ctenophores and cnidarians. *Hydrobiologia* 530-531(1-3):549-556.

15. Tamm SL (2014) Cilia and the life of ctenophores. *Invertebr. Biol.* 133(1):1-46.
16. Hyman LH (1940) *The Invertebrates. Volume 1. Protozoa Through Ctenophora* (McGraw-Hill).

17. Martindale MQ (1986) The ontogeny and maintenance of adult symmetry properties in the ctenophore, Mnemiopsis mccradyi. *Developmental biology* 118(2):556-576.

18. Tamm SL (1984) Mechanical synchronization of ciliary beating within comb plates of ctenophores. *J Exp Biol* 113:401-408.

19. Franc J-M (1978) Organization and Function of Ctenophore Colloblasts: An Ultrastructural Study. *Biological Bulletin* 155(3):527-541.

20. Matsumoto GI & Harbison GR (1993) In situ observations of foraging, feeding, and escape behavior in three orders of oceanic ctenophores: Lobata, Cestida, and Beroida. *Marine Biology* 117(2):279-287.

21. Bumann D & Puls G (1997) The ctenophore Mnemiopsis leidyi has a flow-through system for digestion with three consecutive phases of extracellular digestion. *Physiol Zool* 70(1):1-6.

22. Freeman G & Reynolds GT (1973) The development of bioluminescence in the ctenophore Mnemiopsis leidyi. *Developmental biology* 31(1):61-100.

23. Jager M, Chiori R, Alie A, Dayraud C*, et al.* (2011) New Insights on Ctenophore Neural Anatomy: Immunofluorescence Study in Pleurobrachia pileus (Muller, 1776). *J. Exp. Zool. Part B* 316B(3):171-187.

24. Moroz LL (2015) Convergent evolution of neural systems in ctenophores. *J Exp Biol* 218(Pt 4):598-611.

25. Reverberi G (1974) Experimental embryology of marine and freshwater invertebrates.–587 pp. Amsterdam – London: North Holland Publishing Company 1971. Hfl. 100.—, £ 11.65. *Internationale Revue der gesamten Hydrobiologie und Hydrographie* 59(4):590-590.

26. Tamm SL (2014) Formation of the Statolith in the Ctenophore Mnemiopsis leidyi. *Biol Bull* 227(1):7-18.

27. Chun C (1880) *Die ctenophoren des golfes von Neapel, und der angrenzenden meeres-abschnitte. Eine monographie von Carl Chun* (W. Engelmann, Leipzig).

28. Mortensen T (1915) On regeneration in Ctenophores. *Kjobenhavn Nath Medd* 66:(45.

29. Coonfield BR (1936) Regeneration in *Mnemiopsis Leidyi*, Agassiz. *The Biological Bulletin* 71(3):421-428.

30. Coonfield BR (1937) The Regeneration of Plate Rows in *Mnemiopsis Leidyi,* Agassiz. *Proc Natl Acad Sci U S A* 23(3):152-158.

31. Coonfield BR & Goldin A (1937) The Problem of a Physiological Gradient in *Mnemiopsis* During Regeneration. *The Biological Bulletin* 73(2):197-204.

32. Coonfield BR (1937) Symmetry and Regulation in *Mnemiopsis Leidyi*, Agassiz. *The Biological Bulletin* 72(3):299-310.

33. Freeman G (1967) Studies on regeneration in the creeping ctenophore, *Vallicula multiformis*. *J Morphol* 123(1):71-83.

34. Fischel A (1897) Experimentelle Untersuchungen am Ctenophorenei. *Archiv für Entwickelungsmechanik der Organismen* 6(1):109-130.

35. Farfaglio G (1963) Experiments on the formation of the ciliated plates in Ctenophores. *Acta Embryol. Morphol. Exp* 6:191-203.

36. Martindale MQ & Henry JQ (1996) Development and regeneration of comb plates in the ctenophore Mnemiopsis leidyi. *The Biological Bulletin* 191(2):290.

37. French V, Bryant P, & Bryant S (1976) Pattern regulation in epimorphic fields. *Science* 193(4257):969-981.

38. Henry JQ & Martindale MQ (2000) Regulation and Regeneration in the Ctenophore Mnemiopsis leidyi. *Developmental biology* 227(2):720-733.

39. Andrilenas KK & Moroz L (2010) The Neurogenic Effect of Injury and Regeneration in Ctenophores. *Integr. Comp. Biol.* 50:E200-E200.

40. Castrillon DH, Quade BJ, Wang TY, Quigley C, & Crum CP (2000) The human VASA gene is specifically expressed in the germ cell lineage. *Proc Natl Acad Sci U S A* 97(17):9585-9590.

41. Aravin AA, Hannon GJ, & Brennecke J (2007) The Piwi-piRNA Pathway Provides an Adaptive Defense in the Transposon Arms Race. *Science* 318(5851):761-764.

42. Alie A, Leclere L, Jager M, Dayraud C*, et al.* (2011) Somatic stem cells express Piwi and Vasa genes in an adult ctenophore: ancient association of "germline genes" with stemness. *Developmental biology* 350(1):183-197.

43. Tamm SL (2012) Regeneration of ciliary comb plates in the ctenophore Mnemiopsis leidyi. i. morphology. *J Morphol* 273(1):109-120.

44. Wang Z, Gerstein M, & Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics* 10(1):57-63.

45. Parker TJ & Haswell WA (1897) *A text-book of zoology, by T. Jeffery Parker ... and William A. Haswell ... With illustrations* (Macmillan and co., limited;, London).

46. Partek I (2014) *Partek® Genomics Suite®* (Partek Inc., St. Louis).

47. Petersen TN, Brunak S, von Heijne G, & Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8(10):785-786.

48. Emanuelsson O, Nielsen H, Brunak S, & von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* 300(4):1005-1016.

49. CLC Genomics Workbench 7.0.3 (http://www.clcbio.com).

50. Petersen TN, Brunak S, von Heijne G, & Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Meth* 8(10):785-786.

51. Doron-Mandel E, Fainzilber M, & Terenzio M (2015) Growth control mechanisms in neuronal regeneration. *FEBS Lett* 589(14):1669-1677.

52. Charge SB & Rudnicki MA (2004) Cellular and molecular regulation of muscle regeneration. *Physiol Rev* 84(1):209-238.

53. Letunic I, Doerks T, & Bork P (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res* 43(Database issue):D257-260.

54. Klar A, Baldassare M, & Jessell TM (1992) F-spondin: a gene expressed at high levels in the floor plate encodes a secreted protein that promotes neural cell adhesion and neurite extension. *Cell* 69(1):95-110.

55. Mason PH, Dominguez DJ, Winter B, & Grignolio A (2015) Hidden in plain view: degeneracy in complex systems. *Biosystems* 128:1-8.

56. Lee JH, Kwon EJ, & Kim do H (2013) Calumenin has a role in the alleviation of ER stress in neonatal rat cardiomyocytes. *Biochemical and biophysical research communications* 439(3):327-332.

57. Linde CI, Feng B, Wang JB, & Golovina VA (2013) Histidine triad nucleotide-binding protein 1 (HINT1) regulates Ca(2+) signaling in mouse fibroblasts and neuronal cells via store-operated Ca(2+) entry pathway. *Am J Physiol Cell Physiol* 304(11):C1098-1104.

58. Aravind P, Mishra A, Suman SK, Jobby MK*, et al.* (2009) The betagamma-crystallin superfamily contains a universal motif for binding calcium. *Biochemistry* 48(51):12180-12190.

59. Yang X, Wang Q, Gao Z, Zhou Z*, et al.* (2013) Proprotein convertase furin regulates apoptosis and proliferation of granulosa cells in the rat ovary. *PLoS One* 8(2):e50479.

60. Lin JH, Li H, Yasumura D, Cohen HR*, et al.* (2007) IRE1 signaling affects cell fate during the unfolded protein response. *Science* 318(5852):944-949.

61. Carette JE, Guimaraes CP, Varadarajan M, Park AS*, et al.* (2009) Haploid genetic screens in human cells identify host factors used by pathogens. *Science* 326(5957):1231-1235.

62. Krupczak-Hollis K, Wang X, Dennewitz MB, & Costa RH (2003) Growth hormone stimulates proliferation of old-aged regenerating liver through forkhead box m1b. *Hepatology* 38(6):1552-1562.

63. Maki N, Martinson J, Nishimura O, Tarui H*, et al.* (2010) Expression profiles during dedifferentiation in newt lens regeneration revealed by expressed sequence tags. *Molecular Vision* 16:72-78.

64. Roberson DW, Alosi JA, Mercola M, & Cotanche DA (2002) REST mRNA expression in normal and regenerating avian auditory epithelium. *Hearing Research* 172(1–2):62-72.

65. Cowles MW, Omuro KC, Stanley BN, Quintanilla CG, & Zayas RM (2014) COE Loss-of-Function Analysis Reveals a Genetic Program Underlying Maintenance and Regeneration of the Nervous System in Planarians. *PLoS Genet* 10(10):e1004746.

66. Matsuda H, Yokoyama H, Endo T, Tamura K, & Ide H (2001) An Epidermal Signal Regulates Lmx-1 Expression and Dorsal–Ventral Pattern during Xenopus Limb Regeneration. *Developmental biology* 229(2):351-362.

67. Postic C, Dentin R, Denechaud P-D, & Girard J (2007) ChREBP, a Transcriptional Regulator of Glucose and Lipid Metabolism. *Annual Review of Nutrition* 27(1):179-192.

68. Rodier G, Kirsh O, Baraibar M, Houles T*, et al.* (2015) The transcription factor E4F1 coordinates CHK1-dependent checkpoint and mitochondrial functions. *Cell Rep* 11(2):220-233.

69. Ota M & Sasaki H (2008) Mammalian Tead proteins regulate cell proliferation and contact inhibition as transcriptional mediators of Hippo signaling. *Development* 135(24):4059-4069.

70. Martoglio B & Dobberstein B (1998) Signal sequences: more than just greasy peptides. *Trends Cell Biol* 8(10):410-415.

71. Kundaje A, Meuleman W, Ernst J, Bilenky M*, et al.* (2015) Integrative analysis of 111 reference human epigenomes. *Nature* 518(7539):317-330.

72. Lyko F, Foret S, Kucharski R, Wolf S*, et al.* (2010) The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol* 8(11):e1000506.

73. Lister R, Mukamel E A, Nery JR, Urich M*, et al.* (2013) Global epigenomic reconfiguration during mammalian brain development. *Science* 341(6146):1237905.

74. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312-1313.

75. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792-1797.

76. Sun D, Xi Y, Rodriguez B, Park H J*, et al.* (2014) MOABS: model based analysis of bisulfite sequencing data. *Genome Biol* 15(2):R38.

BIOGRAPHICAL SKETCH

Rachel Sanford received her Bachelors of Science in zoology and psychology from the University of Florida in 2011. She graduated *cum laude* and made the Dean's List from 2008 to 2011. From there she participated in a research apprenticeship in the spring of 2012 at University of Washington's Friday Harbor Laboratory and joined the lab of Dr. Leonid Moroz at the University of Florida's Whitney Laboratory for Marine Biosciences in 2013. She graduated with a Master of Science degree majoring in medical science in 2015.