Spring 2008

Characterization of Sox family members in sponge stem cells and during development

Karen E. Leeds
University of Richmond

Characterization of *Sox* Family Members in Sponge Stem Cells and during Development

Karen E. Leeds

Honors Thesis

in

Department of Biology
University of Richmond
Richmond, VA

April 25. 2008

Advised by Dr. April L. Hill

ABSTRACT

Sponges are considered the oldest and most basal part of the metazoan lineage and therefore possess a unique set of gene families that are highly conserved among all animals. One of these gene families is known as *Sox* transcription factors. *Sox* genes are known to play important roles in complex animals such as the specification of the primary layers of the body, determination of sex, and most recently induction of induced pluirpotent stem cells (iPS cells) from both human and mouse fibroblasts with the help of three other transcription factors. We have found that two different demosponge species *Halichondria bowerbanki* and *Ephydatia muelleri* already possess at least three distinct *Sox* transcription factors using evolutionary PCR and bioinformatic approaches. This supports the hypothesis that eumetazoans possessed at least three distinct *Sox* genes. To better understand the role that each of these *Sox* transcriptions may have played during the evolution and development of complex body plans we conducted both temporal and spatial expression analysis. We performed both RT-PCR and Real Time RT-PCR on all *Sox* genes isolated from both marine and freshwater sponges. These studies demonstrated that *Sox* genes are expressed during larval and adult development in marine sponges and during the development of gemmules to fully functioning adult stages in freshwater sponges. In situ hybridization studies have revealed that one of the *Sox* genes is expressed in developing embryos of the marine sponge. Furthermore, at least one *Sox* gene in both the marine and freshwater sponges shows distinct expression in archeocytes, the sponge stem cell. The presence of *Sox* expression in archeocytes suggests a role in cell specification or differentiation, as well as suggests a possible role in evolution of multicellularity. Therefore this study provides an establishment for future studies aimed at testing the hypothesis that the *Sox* transcription factor family had a crucial role in the evolution of animal multicellularity.

INTRODUCTION

Recent studies demonstrate that both human and mouse fibroblasts can be reprogrammed in vitro into induced pluripotent stem cells (iPS cells) through retroviral transduction of the same four transcription factors *Sox*2, *c-Myc*, *Oct*3/4, and *Klf*4 (Takashashi et al, 2007; Yu et al, 2007; Park et al, 2008). These iPS cells are similar to human embryonic stem cells in a variety of ways including morphology and gene expression. Furthermore, the iPS cells are comparable to embryonic stem cells because these cells can differentiate into cell types from all three germ layers, the mesoderm, ectoderm, and endoderm (Takashashi et al, 2007). Although human iPS cells are important for understanding human diseases and developing methods of treatment, this research can also lead us to ask questions about how multicellularity evolved in the metazoan branch of the tree of life and elucidate the role of the development of undifferentiated cells into a fully functional adult with many distinct cell types (Mueller, 2006).

One way to study the question of multicellularity and the evolution of the metazoan is by looking at the oldest and most living animals, the porifera. Sponges are an excellent model organism, especially for evolutionary studies in development because they evolved at least 580 million years ago. They can be considered a living fossil because they represent a basal metazoan phylum that existed prior to the Cambrian explosion 542 millions years ago. The Cambrian explosion was a time period when most of the complex animal body plans appeared in the fossil record (Carroll et al, 2005). The sponges were in existence before this occurred and are the most ancient and basal member of the complex animal lineage. It is because the sponge is basal and less complex in body plan they can be used as a model to help reveal the genetic and molecular events that may have occurred from the transition from one-celled eukaryotes to complex animals (Cetkovic and Lukic-Bilela, 2003). In order to study the evolution of multicellularity and

the differentiation of one cell into many types of cells I focused on one of the four transcription factor families that induce pluripotent stem cells, *Sox*, to see if sponges already possessed members of this transcription factor family.

The *Sox* transcription factor family contains a large number of genes that possess the 79-amino acid DNA binding domain known as the High Motility Group or the HMG Box (Figure 1). The *Sox* family is divided into eight groups (A-H) based on their primary sequence and structure and have been found in a variety of metazoans including approximately 30 vertebrate genes and 14 invertebrate genes. Furthermore, the mouse and human genomes contain 20 orthologous pairs of *Sox* genes (Jager et al, 2006; Magie et al, 2005; Schepers et al, 2002). *Sox* proteins are known for a variety of roles in the metazoan lineage especially in developmental processes, including neural crest specification, gastrulation, the development of the central nervous system, and chondrogenesis (Magie et al, 2005; Schepers et al, 2002). These transcription factors are also known to play roles in cell differentiation, cell specification, and germ layer formation (Jager et al, 2006). Furthermore, the *Sox* family is also commonly known to be related to the mammalian SRY genes, or the sex-determining genes (Collignon et al, 1996; Magie et al, 2005).

Recent studies indicate the presence of *Sox* genes in the basal lineages of the metazoans including 14 *Sox* genes in the cnidarian, *Nematostella vectensis*, 3 *Sox* genes in *Amphimedon queenslandica*, a type of marine demosponge found only on the Great Barrier Reef in Australia, and 3 *Sox* genes in *Ephydatia muelleri*, a freshwater demosponge. (Larroux et al, 2006; Jager et al, 2006, Magie et al 2005) These were the first studies that revealed the presence of *Sox* transcription factors in basal metazoans. It was also recently reported that there is one *Sox* family member also present in the genome of the choanoflagellate, *Monosiga brevicollis*, the most

common ancestor to the metazoans (King et al, 2008). In the cnidarian, *Nematostella vectensis*, spatial expression analysis studies indicate *Sox* is present in early polyp stages, ectodermal cells, and the endodermal lining of the pharynx, implicating roles of neural cell differentiation and germ layer specification (Magie et al, 2005). However, to this date no spatial or developmental expression has been reported on the *Sox* genes in sponges.

Therefore, based on these foundations, this study has two main focuses. First, to determine the minimal number of *Sox* genes present in the demosponge genome through bioinformatic and PCR-based approaches, using two different species *Ephydatia muelleri* and *Halichondria bowerbanki*, a marine demosponge. Second, I also want to characterize the expression patterns of *Sox* genes using both temporal and spatial expression analysis. My hypothesis is that the common ancestor to the demosponges possessed 3-4 *Sox* gene family members and that at least one of these genes is involved in stem cell determination and/or specification in sponges. These studies will also help us to learn more about all the roles *Sox* genes play in the basal metazoans. Therefore this study will give us further insight into the evolution of multicellularity and the first roles that these transcription factor families played in the evolution of animals.

MATERIALS AND METHODS

Collection of Sponges

Halichondria bowerbanki was collected during the adult and larval stages from the Cheseapeake Bay at Virginia Institute of Marine Science, Gloucester Point, Virginia. *Ephydatia muelleri* was collected by Sally Leys in Canada. Tissue was processed at designated developmental stages and stored at –80^{o}C.

Isolation of *Sox* sequences

Halichondria bowerkbanki and *Ephydatia muelleri* genomic DNA was isolated using the CTAB method and mRNA from either free-swimming larvae or aggregated adult tissue was isolated using the RNAeasy Kit (Qiagen) according to manufacture's protocol. cDNA was formed using Thermoscript Reverse Transcriptase (Invitrogen) using an oligo(dT) primer.

For *EmSox2* gene isolation in *Ephydatia muelleri* intial degenerate PCR primers were designed using the *Sox* gene sequences from *Amphimedon queenslandica* through a local basic alignment search tool (BLAST) using the trace archives of the *Amphimedon queenslandica* genome provided by the Joint Genome Institute. The amino acid sequence of *Nematostella vectensis* Sox family protein 1 was intially used in a tblastn search against the trace archives of *Amphimedon queenslandica*. The retrieved nucleotide sequences were then confirmed and translated into protein sequences to design the degenerate primers. The primer combination VKRPMN (foward) and (YKYKPKR) reverse was designed for isolation of the *Sox* conserved domain and produced a band of the expected size (201 bp). The PCR conditions that yieleded these results were 2 min at 95 ^{o}C followed by 30s at 95 ^{o}C, 30s at 40 ^{o}C, 1 min at 72 ^{o}C for 35 cycles, followed by 5 min at 72 ^{o}C.

For *HbSox3* gene isolation in *Halichondria bowerbanki* intial degenerate PCR reactions were performed using the same degenerate primers as described above again for the isolation of the *Sox* conserved domain and produced a band of the expected size (201 bp). The PCR conditions that yielded these results were 2 min at 95 ^{o}C followed by 30s at 95 ^{o}C, 30s at 40 ^{o}C, 1 min at 72 ^{o}C for 35 cycles, followed by 5 min at 72 ^{o}C.

For the *EmSox1 and EmSox3* gene isolation in *Ephydatia muelleri* initial gene specific PCR reactions were performed using nucleotide sequenced described in Jager et al. (2006) for isolation of *EmSox1* and *EmSox3* family members from the sponge *Ephydatia muelleri*. The primer combination used was 5'-TCAATCGAGCGCAAGAAG-3' (forward) and 5'-CCTGGAGAGAACACCCGGA-3' (reverse) were designed for *EmSox1* class sequences and produced a band of the expected size (162 bp). The primer combination used was 5'-GCTGGAACGGAGGAAGATGA-3' (foward) and 5'-CACATGAAGCAGTACCCCG-3' (reverse) were designed for *EmSox3* class sequences and produced a band of the expected size (159 bp). The PCR conditions that yielded these results were 2 min at 94^{o}C followed by 30s at 94^{o}C, 30s at 58^{o}C, 1 min at 72^{o}C for 30 cycles, followed by 5 min at 72 ^{o}C.

All resulting PCR products described above were visualized via agarose gel electrophoresis and bands of expected sizes were excised and cloned using the TOPO TA Cloning Kit (Invitrogen). Clones were sequenced using the SequiTherm EXCEL II kit (Epicenter) on a LiCor DNA Sequencing System.

Isolation of 3' ends of Sox genes

Halichondria bowerkbanki and *Ephydatia muelleri* 3' RACE pools were created according to Clontech's SMART RACE 3' cDNA Isolation Protocol from *Halichondria* larvae and *Ephydatia* tissue. The Clontech procedure was followed using the AUAP primer provided for the poly-A tail in combination with gene specific primers for each of the *Sox* genes. The primer for *EmSox1* was 5'-TCAATCGAGCGCAAGAAG-3' followed by a nesting with the primer 5'-GCCGAGAGAGAACCGCGCTTG-3'. The primer for *EmSox2* was 5'-ACGCCTTCATGGTGTGGTCT-3' followed by a nesting with the primer 5'-CGGTTGGGGGCGCAGTGGAAAG-3'. The primer for *EmSox3* was 5'-

GCTGGAAAGGAGGAAGATGA -3' followed by a nesting with the primer 5'-TGGGCAAGCTGTGGCGGTTACTC-3'. The reactions were performed under the following PCR conditions: 3 min at 94°C, followed by 30s at 94°C, 30s at 58°C, 1 min at 72°C- for 35 cycles, followed by 5 min at 72°C. The primer for *HbSox3* was 5'-CGCATTTATGGTCTGGGCTCA -3' followed by a nesting with the primer 5'-AACTGGAGAGGAGAAAGATG-3'. The reactions were performed under the following PCR conditions: 3 min at 94°C, followed by 30s at 94°C, 30s at 60°C, 1 min at 72°C- for 35 cycles, followed by 5 min at 72°C. All resulting PCR products described above were visualized via agarose gel electrophoresis. A Southern blot was performed in order to determine which hybridized bands were the 3' end of both *EmSox* and *HbSox* sequences. These bands were excised and cloned using the TOPO TA Cloning Kit (Invitrogen). Plasmids were sent to Virginia Commonwealth University for sequencing in both directions.

RNA Isolation, RT-PCR, and Real Time RT-PCR

Halichondria bowerbanki larvae were collected by placing individual sponges in beakers of sterile, filtered seawater. Reproductive sponges released larvae into the water column and newly released larvae were collected by pipetting. Larvae were cultured in 24-well plates in filtered, sterile seawater, which was replaced daily. Tissue was collected from five developmental stages of *Halichondria*: free swimming larvae (Free Swim), larvae skating across the surface of the plate (Skate), larvae attached to the plate surface (Attached), larvae that had begun to grow and spread across the plate surface (Spread), and small adults that had formed a water pumping chamber (Rhagon). Adult sponge tissue was also collected. Tissues were stored in RNAlater (Ambion) overnight and then placed at -80^{o}C for subsequent RNA isolation.

Ephydatia muelleri gemmules were obtained and collected from six development stages of *Ephydatia*: resting gemmule, cells in stasis (Stage 0), growth begins and cells migrate out from the gemmule coat (Stage 1), archeocytes begin to differentiate cell types (Stage 2), choanocyte chambers form, sponge begins to filter water (Stage 3), oscule forms, sponge fully functional (Stage 4), sponge grows in size (Stage 5). Tissues were stored in RNAlater (Ambion) overnight and then placed at -80^{o}C for subsequent RNA isolation.

Halichondria bowerbanki and *Ephydatia muelleri* RNA was isolted using the RNAeasy Kit (Qiagen) and treated with DNase I to eliminate contaminating genomic DNA. For RT reactions, 500 ng of RNA was reverse transcribed using the Thermoscript RT kit (Invitrogen) and subsequent PCR reactions were carried out using Platinum Taq DNA polymerase (Invitrogen). To amplify *HbSox3* from *Halichondria bowerbanki* the primers used were forward: 5‘-CGCATTTATGGTCTGGGCTCA-3‘ and reverse: 5‘-AACACATGAAGCAATACCCCG-3‘ under PCR conditions: 94^{o}C for 3 min followed by 15s at 94^{o}C, 25s at 60^{o}C, 1 min at 72^{o}C for 30 cycles and then 2 min at 72^{o}C. To amplify *EmSox2* from *Ephydatia muelleri* the primers used were forward 5‘-ACGCCTTCATGGTGTGGTCT-3‘ and reverse: 5‘-TACATCGAGGAGGCCAAGCG-3‘ under PCR conditions: 94^{o}C for 3 min followed by 15s at 94^{o}C, 25s at 58^{o}C, 1 min at 72^{o}C for 30 cycles and then 2 min at 72^{o}C. RT-PCR positive controls reactions were performed using actin gene primer sequences (as described in Hill et al., 2004).

Real Time RT-PCR was performed using RT reactions and stages described above and subsequent PCR reactions were carried out using SYBR Green ER (Invitrogen). *Ephydatia muelleri* and *Halichondria bowerbanki* RT stages were used in a 1:2 dilution for *EmSox2* and *EmSox3*, and 1:1 dilution for *EmSox1* and *HbSox3*. The same primer reactions and PCR

conditions for *HbSox3* and *EmSox2* were used as described above. PCR amplication of three dilutions of the cDNA of each of these genes (*HbSox*3: 41.4 pg/uL, 4.14 pg/uL, 0.414 pg/uL and *EmSox*2: 25 pg/uL, 2.5 pg/uL, and 0.25 pg/uL) were used to generate a standard curve. To amplify *EmSox1* from *Ephydatia muelleri* three dilutions of the cDNA were also used (27.6 pg/uL, 2.76 pg/uL, 0.276 pg/uL). The primers used were forward 5‘-TCAATCGAGCGCAAGAAG-3‘ and reverse: 5‘-CCTGGAGGAACACCCGGA-3‘. To amplify *EmSox3* from *Ephydatia muelleri* three dilutions of the cDNA were used (19.5 pg/uL, 1.95 pg/uL, 0.195 pg/uL). The primers used were forward 5‘-GCTGGAAAGGAGGAAGATGA-3‘ and reverse: 5‘-CACATGAAGCAGTACCCCG-3‘. Both *EmSox1* and *EmSox3* reactions were performed under PCR conditions described above from *EmSox2*. RT-PCR positive controls reactions were performed using actin gene primer sequences from *Halichondria* and elongation factor-1 from *Ephydatia*.

In situ hybridization

Halichondria bowerbanki tissues (adult, reproductive adult, larvae) were fixed overnight in 4% paraformaldehyde, 0.035% glutaraldehyde in 1X PBS and then transferred into ascending concentrations of ethanol, and stored in 100% ethanol at –80^{o}C. In situ protocol was developed from several labs including Nipam Patel for *Parhyale* in situ hybridizations and Scott Nichols’s RNA in situ hybridization protocol on paraffin embedded tissues were both used with some modifications. Fixed tissues were rehydrated through two xylene washes and an ethanol and DEPC water series. Tisses were prepared for prehybridization through the following series of washes: one wash of 1X PBS, two washes of 1X PBS with 100 mM glycine, one wash 1x PBS containing 0.3% Triton X-100, two washes 1X PBS, one wash of 1X TE (pH 8) containing Proteinase K (1 ug/mL) at 37^{o}C. The tissues were then re-fixed in 4% paraformaldehyde in PBS.

Subsequently, tissues were washed in two washes of 0.1 TEA buffer (pH 8) containing 0.25% acetic anhydride, followed by two washes of 1X PBS. Tissue was then prehybridized in hybridization buffer (50% formamide, 5X SSC, 50 μg/ml heparin, 0.25% Tween-20, 1% SDS, 100 μg/ml single-stranded DNA; pH 5) for one hour at 42 °C. All probes were labeled using the Dig RNA labeling kit (Roche®). For *HbSox3*, the region between foward: 5‘-CGCATTTATGGTCTGGGCTCA-3‘ and reverse: 5‘-AACTGGAGAGGAGAAAGATGA-3‘ was cloned and used as a riboprobe. All products were cloned in the TOPO TA Dual Promoter Cloning Vector (Invitrogen) prior to probe preparation and both sense and antisense products were generated for each probe. After overnight hybridization at 42°C tissue was washed in three washes of 2X SSC, two washes of 1X SSC, one wash of NTE buffer containing 20 ug/mL RNAse A at 37 °C , and followed by two washes of 0.1X SSC at 37 °C. After tissue was incubated in Buffer 1 (100 mM Tris-HCl (pH 7.5), 10 mM EDTA), followed by Buffer 1 containing 0.1% Triton X-100, 2% normal sheep serum to block nonspecific binding of antibody and then processed for staining. In cases where sections are shown, tissue was processed through alcohols and xylene, embedded in parrafin wax, and cut on a Leica Microsystems RM2245 rotary microtome.

Ephydatia muelleri tissues were fixed overnight in 4% paraformaldehyde, 0.035% glutaraldehyde in 1/4 HS (Holtfrer Solution) and then transferred into ascending concentrations of ethanol, and stored in 100% ethanol at –80°C. In situ protocol was adapted from Funayama et al and CSH with some modifications. Fixed tissues were rehydrated through an ethanol and 1/4 HS series, followed by one wash of PTw (1X PBS containing 0.1% Tween-20), subsequently re-fixed in 4% paraformaldehyde in 1X PBS at 4 °C, and washed twice with PTw. Tissue was then prehybridized in hybridization solution (50% formamide, 5X SSC, 50 μg/ml heparin, 0.25%

Tween-20, 1% SDS, 10 mM DTT, 100 μg/ml single-stranded DNA; pH 5) at 50 °C for 2 hours. All probes were labeled using the Dig RNA labeling kit (Roche®). For *EmSox2*, the region between : 5‘-ACGCCTTCATGGTGTGGTCT-3‘ and reverse: 5‘-TACATCGAGGAGGCCAAGCG-3‘ was cloned and used as a riboprobe. All products were cloned in the TOPO TA Dual Promoter Cloning Vector (Invitrogen) prior to probe preparation and both sense and antisense products were generated for each probe. After overnight hybridization at 50°C tissue was washed 7 times in hybridization solution at 50°C and gradually processed to room temperature through half washes in 1/4 HS and hybridization solution. After a few washes in Maleic acid buffer 1 (0.1M maleic acid, 0.15M NaCl, 0.1% Tween-20) tissue was incubated in Maleic acid buffer 1 containing 1% BSA to block nonspecific binding of antibody and then processed for staining.

RESULTS

How many *Sox* genes are present in demosponges?

One distinct *Sox* gene (*HbSox3*) so far has been isolated from the marine sponge *Halichondria bowerbanki* using degenerate PCR (Figure 2). This particular Sox gene is most similar to the *Amphimedon queenslandica Sox* C-like gene on the amino acid level when compared to Sox sequences in GenBank in NCBI. Currently, we have not yet isolated other *Sox* genes from *Halichondria*; however, this does indicate other *Sox* genes are not present. Further work needs to be done by developing new degenerate PCR primers and using other stages of development to search for additional *Sox* genes in this particular demosponge.

Three distinct *Sox* genes were isolated from the freshwater sponge *Ephydatia muelleri* using both degenerate PCR and gene specific primers adapted from Jager et al, (2006). Both

EmSox2 and *EmSox3* sequences obtained were identical on the amino acid level to the partial *Sox* sequences obtained by Jager et al, (2006) in *Ephydatia muelleri* (Figure 4A-B; Figure 5A-B). Furthermore, we also obtained part of the 3' end of *EmSox3* using 3' RACE (Figure 5C-D). The *EmSox1* sequence was also identical on the amino acid level to the sequence obtained by Jager et al, (2006) (Figure 3A, 3C). However, we also found an alterative form of the gene, possibly caused by alternative splicing (Figure 3B). Each of the three *Sox* genes are similar to a particular *Sox* gene in *Amphimedon queenslandica*, as stated above on the amino acid level. *EmSox1* is most similar to *AqSoxF*, *EmSox2* is most similar to *AqSoxB*, and *EmSox3* is most similar to *AqSoxC*.

What is the spatial and temporal expression pattern of each of the *Sox* genes in sponges?

We used both Reverse Transcriptase-PCR and Real Time Reverse Transcriptase-PCR to determine the stages of development *HbSox3* is expressed. *HbSox3* is expressed in all stages of development; however it most highly expressed in the free swimming larval stages and larval attachment stage (Figure 6; Figure 7).

In order to determine where *HbSox3* is expressed in both developing and adult tissue in situ hybridization was performed. *HbSox3* is expressed in both early stage embryos (Figure 8A-B; Figure 8E-F) and pre-release larvae (Figure 8C-F) of the reproducing adult sponge. Furthermore, *HbSox3* appears to be expressed in the archeocyte cells or the putative stem cells of the adult sponge (Figure 9A-D).

In *Ephydatia*, Real Time Reverse Transcriptase-PCR was completed on both *EmSox1* and *EmSox3* to determine the stages of development these genes are present during metamorphosis from gemmule to hatched adult. Both *EmSox1* and *EmSox3* are present in all stages of

development however *EmSox1* is most highly expressed at stage 0 (Figure 11) and *EmSox3* is most highly expressed at stage 1 (Figure 12).

We also performed Reverse Transcriptase-PCR and Real Time Reverse Transcriptase-PCR to determine the stages of development *EmSox2* is expressed. *EmSox2* is expressed in all stages of development; however it is also most highly expressed at stage 0 (Figure 13; Figure 14).

In order to determine where *EmSox2* is expressed in the developing sponge in situ hybridization was also performed. *EmSox2* is expressed in one distinct group of cells surrounding the developing canal system of the sponge (Figure 15A, 15C). These specific cells are believed to be archeocyte cells (Figure 15B) that are present in the regions that will develop into the feeding and pumping chambers (Figure 15E-H). However, *EmSox2* expression is not present in the choanocyte chambers themselves (Figure 15D).

DISCUSSION

Eumetazoans contained at least three *Sox* genes

Sponges contain a unique set of genes because of their basal position in the animal lineage. *Sox* transcription factors are a part of this group of genes that play roles in the development not only in sponges but the entire metazoan lineage. In this particular study, through bioinformatics and evolutionary PCR approaches, we were able to demonstrate the presence of at least three *Sox* genes from the freshwater sponge *Ephydatia muelleri*, and the presence of at least one *Sox* gene from *Halichondria bowerbanki*. These results verify the presence of at least three *Sox* genes from two Demospongiae class sponges, including *Ephydatia muelleri* and *Amphimedon queenslandica*. Three *Sox* genes also have been found in the

Calcispongia class of sponge (Larroux et al, 2006; Jager et al, 2006). The presence of at least three *Sox* genes in sponges suggests that the *Sox* transcription factors duplicated and diversified early in the animal lineage. Therefore this implies the common ancestor of eumetazoans already had a set of three *Sox* genes. Additionally, the presence of these three ancestral *Sox* genes in sponges suggests that it was these three *Sox* gene families throughout animal evolution duplicated and diversified into the eight different subclasses of *Sox* genes present in the metazoan lineage today.

***HbSox*3 expression suggest roles in embryonic and larval stages of development**

We have demonstrated using in situ hybridization, expression of *HbSox*3 shown in both developing embryos, as well as pre-release larvae. Therefore our preliminary evidence suggests that *Sox* may play a role in the early development of sponge embryos. However, further studies need to be done in order to determine the specific role *HbSox*3 is playing in the development of the embryo since the expression of *HbSox*3 was throughout the early embryo and not only in a particular cell type. It may play a more generalized role in early cell proliferation or communication.

Furthermore, we also have demonstrated the possible role of *HbSox*3 involvement in larval stages of *Halichondria* from both RT-PCR and Real-Time RT-PCR. The high expression at the free swim stage and attached stage could suggest further roles in the development of the larvae at the free swim stage and a possible role in the differentiation of sponge cells as the larvae grows across its substrate at the attached stage, before it develops into a rhagon. However, further studies such as in situ hybridization at each larval stage needs to be performed to help determine the specific role of *HbSox*3 at each stage of larval development.

*EmSox*1-3 expression during freshwater sponge metamorphosis

We also have established the possible role of *Sox* involvement in the gemmule hatching stages of *Ephydatia* from both RT-PCR and Real-Time RT-PCR. *EmSox*1-3 expression is seen in every stage of *Ephydatia* development and metamorphosis from gemmule to adult; however each gene is turned on at different levels during each stage. This suggests that each *EmSox* gene is playing a distinct role in the development of the *Ephydatia* tissue. However, further studies are necessary to understand the specific role each *Sox* gene is playing at every stage. *Ephydatia* stages 1-5 gemmule hatching and metamorphosis most closely resemble *Halichondria* larval attachment through rhagon stages. Future experiments should closely evaluate *Halichondria* and *Ephydatia* orthologs during this developmental time frame to compare if roles and patterns are possible conserved.

*HbSox*3 and *EmSox*2 expression in adult sponge tissue suggest role in archeocyte differentiation

We have established by using in situ hybridization, expression of both *HbSox3* in *Halichondria* and *EmSox2* in *Ephydatia* shown in archeocyte cells surrounding the choanocyte chambers. Archeocytes cells are putative stem cells of the sponge that differentiate into at least ten cell types (Funayama et al, 2005). Our preliminary evidence implies the expression of *Sox* orthologs in the archeocyte cells of the sponge and suggests that *Sox* genes may play roles in cell differentiation and specification.

Sox transcription factors and the evolution of multicellularity

We have demonstrated that at least two *Sox* orthologs *HbSox*3 and *EmSox*2, are expressed in archeocyte cells or putative sponge stem cells, which as mentioned above suggests a possible role in cell differentiation. A possible role of *Sox* genes in sponges in cell differentiation would

imply that *Sox* may have had a role in the evolution of multicellularity. This is further supported in the recent report of one *Sox* family member in the genome of the choanoflagellate, *Monosiga brevicollis*. This data demonstrates that *Sox* is one of the oldest transcription factor families in animal genomes. Therefore *Sox* could be involved in regulating gene expression to allow more than a single cell type to exist in one organism (King et al, 2008).

Further evidence to support this hypothesis of *Sox* gene involvement in the evolution of multicellularity is based on recent evidence involving stem cells. It was recently reported that a *Sox* ortholog is one of the four transcription factors that are able to induce pluripotent stem cells from both human and mouse fibroblasts. In one particular study they aimed to induce pluripotent stem cells without the addition of a *Sox* transcription factor. They discovered the cell had embryonic stem-like morphology but it was nullipotent and could not differentiate into any other cells (Yamanaka, 2008). This study further supports *Sox* transcription factors having a critical role in the evolution of multicellularity because in order to get multiple cell types in a single organism, a signal must be sent to tell one cell (such as a stem cell) to differentiate into multiple types of cells.

However, at this point our results do not indicate whether or not the *Sox* transcription factor family had a crucial role in the evolution of multicellularity. Additional studies need to be conducted in both the sponge and choanoflagellate to further test this hypothesis.

This study has demonstrated that the common ancestor to all metazoans possessed at least three distinct *Sox* gene family members. The expression of *Sox* genes in developing embryos, larval, and gemmule stages suggests that *Sox* plays a critical role in the development stages of the adult sponge tissue. Furthermore, the expression of *Sox* in sponge archeocyte cells and its

possible role in cell differentiation can help lead to further studies on testing the hypothesis of *Sox* transcription factors having a crucial role in the evolution of animal multicellularity.

ACKNOWLEDGEMENTS

Technical and Experimental Support:
University of Richmond – Dr. Malcolm Hill, Lisa Warner, Marissa White, Kay Holstien, Aimee Janesky, Kaelyn Krook, and the rest of the Hill lab

University of Alberta – Dr. Sally Leys

Funding:
Summer 2007. Robert F. Smart Award in Biology: "Metazoan germ layer specification genes in sponges: Identification and Expression of Sox Orthologs."

REFERENCES

Carroll, S.B., Grenier, J.K., Weatherbee, S.D. (2005). From DNA to Diversity Molecular Genetics and the Evolution of Animal Design. Malden: Blackwell Publishing. 57-58.

Cetkovic, H. and Lukic-Bilela, L**.** (2003). HMGB2 protein from the marine sponge *Suberites domuncula*. *Food Technology and Biotechnology* **41**, 361-365.

Collignon, J., Sockanathan, S., Hacker, A., Cohen, Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P. N. and Lovell Badge, R. (1996). A comparison of the properties of *Sox*-3 with Sry and two related genes, *Sox*-1 and *Sox*-2. *Development* **122**, 509-520.

Funayama, N., Nakatsukasa, M., Hayashi, T. and Agata, K. (2005). Isolation of the choanocyte in the fresh water sponge, *Ephydatia fluviatilis* and its lineage marker, Ef annexin. *Development Growth & Differentiation* **47**, 243-253.

Hill, A., Tetrault, J., Hill, M. (2004). Isolation and expression analysis of a poriferan Antp-class *Bar*-/*Bsh*-like homeobox gene. *Development Genes and Evolution* **214**, 515-523.

Jager, M., Queinnec, E., Houliston, E. and Manuel, M. (2006). Expansion of the *SOX* gene family predated the emergence of the Bilateria. *Molecular Phylogenetics and Evolution* **39**, 468-477.

King, N., Westbrook, M. J., Young, S. L., Kuo, A., Abedin, M., Chapman, J., Fairclough, S., Hellsten, U., Isogai, Y., Letunic, I. et al. (2008). The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature* **451**, 783-788.

Larroux, C., Fahey, B., Liubicich, D., Hinman, V. F., Gauthier, M., Gongora, M., Green, K., Worheide, G., Leys, S. P. and Degnan, B. M. (2006). Developmental expression of transcription factor genes in a demosponge: insights into the origin of metazoan multicellularity. *Evolution & Development* **8**, 150-173.

Magie, C. R., Pang, K. and Martindale, M. Q. (2005). Genomic inventory and expression of *Sox* and Fox genes in the cnidarian *Nematostella vectensis*. *Development Genes and Evolution* **215**, 618-630.

Muller, W. E. G. (2006). The stem cell concept in sponges (Porifera): Metazoan traits. *Seminars in Cell & Developmental Biology* **17**, 481-491.

Park, I. H., Zhao, R., West, J. A., Yabuuchi, A., Huo, H. G., Ince, T. A., Lerou, P. H., Lensch, M. W. and Daley, G. Q. (2008). Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141-146.

Schepers, G. E., Teasdale, R. D. and Koopman, P. (2002). Twenty pairs of *Sox*: Extent,

homology, and nomenclature of the mouse and human *Sox* transcription factor gene families. *Developmental Cell* **3**, 167-170.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K. and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872.

Yamanaka, S. (2008). Induction of pluripotent stem cells from mouse fibroblasts by four transcription factors. *Cell Proliferation* **41**, 51-56.

Yu, J. Y., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., Slunkin, I., Thomson, J. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917-1920.

FIGURE LEGENDS

Figure 1

Schematic of the *Sox* transcription factor with the HMG-*Sox* Domain.

Figure 2

(A) Nucleotide sequence of *Halichondria bowerbanki Sox3* and (B) the corresponding amino acid sequence.

Figure 3

(A) Nucleotide sequence of *Epyhdatia muelleri Sox1*, (B1) the corresponding amino acid sequence, and (B2) the alternate amino acid sequence possibly caused by alternative splicing in the *EmSox1* gene.

Figure 4

(A) Nucleotide sequence of *Epyhdatia muelleri Sox2* and (B) the corresponding amino acid sequence.

Figure 5

(A) Nucleotide sequence of *Epyhdatia muelleri Sox3* and (B) the corresponding amino acid sequence obtained from gene specific primers designed from Jager at el, 2006). (C) Nucleotide sequence of *Epyhdatia muelleri Sox3* 3' end and (D) the corresponding amino acid sequence obtained by 3' RACE.

Figure 6

RT-PCR expression analysis of *HbSox3* in the stages of development of *Halichondria*. Actin expression is shown as a reference control for expression that is similar in all stages of development.

Figure 7

Real Time RT-PCR expression analysis of *HbSox3* in the stages of development of *Halichondria*. All stages of development are standardized to the control gene actin.

Figure 8

In situ hybridization to paraffin sectioned *Halichondria* adult tissue with *HbSox3* riboprobe. (A) Cross section (30 uM) of *Halichondria* adult tissue with embryos using *HbSox3* antisense probe, staining in the early stage embryos (red arrow). (B) DAPI stain showing the nuclei of each cell (white arrow) corresponding to the tissue in (A). (C) Cross section (30 uM) of *Halichondria* adult tissue with embryos using *HbSox3* antisense probe, staining in the pre-release larvae (black arrow), (D) DAPI stain showing the nuclei of each cell (white arrow) corresponding to the tissue in (C). (E) Cross section (30 uM) of *Halichondria* adult tissue with embryos using *HbSox3* antisense probe, staining in both early stage embryos (red arrow) and pre-release larvae (black arrrow), (F) DAPI stain showing the nuclei of each cell (white arrow) corresponding to the tissue in (E). All sections are shown at 100X.

Figure 9

In situ hybridization to paraffin sectioned *Halichondria* adult tissue with *HbSox3* riboprobe. (A, C) Cross section (30 uM) of *Halichondria* adult tissue using *HbSox3* antisense probe, staining in the archeocyte cells (red arrows). (B, D) DAPI stain showing the nuclei of each cell (white arrows) corresponding to the tissue in (A, C). Sections (A-B) are shown at 200X, and sections (C-D) are shown at 100X.

Figure 10

Epyhdatia muelleri stages of development adapted from Funayama et al, (2005).

Figure 11

Real Time RT-PCR expression analysis of *EmSox1* in the stages of development of *Ephydatia*. All stages of development are standardized to the control gene elongation factor-1.

Figure 12

Real Time RT-PCR expression analysis of *EmSox3* in the stages of development of *Ephydatia*. All stages of development are standardized to the control gene elongation factor-1.

Figure 13

RT-PCR expression analysis of *EmSox2* in the stages of development of *Ephydatia.* Actin expression is shown as a reference control for expression that is similar in all stages of development.

Figure 14

Real Time RT-PCR expression analysis of *EmSox2* in the stages of development of *Ephydatia*. All stages of development are standardized to the control gene elongation factor-1.

Figure 15

Whole mount in situ hybridization to *Ephydatia* tissue with *EmSox2* riboprobe. (A,C) *Ephydatia* tissue using *EmSox2* antisense probe, staining in cells around choanocyte chambers. (B) Cross section of *Ephydatia* tissue stained with vital dye to show cell types, archeocytes are indicated with black arrows. (Photo courtesy of Dr. Sally Leys) (D) DAPI stain showing where choanocyte chambers are present. (E, G) Cross section (30 uM) of *Ephydatia* tissue using *EmSox*2 antisense probe, staining in the archeocyte cells

(red arrows).). (F, H) DAPI stain showing the nuclei of each cell (white arrows) corresponding to the tissue in (E, G). Section (A) is shown at 10X, sections (C-D) are shown at 100X, sections (E-F) are shown at 200X, and sections (G-H) are shown at 400X.

Figure 1

Figure 2

***Halichondria bowerbanki Sox*3**

A: Nucleotide sequence:

5′GTGAAGCGCCCGATGAACGCATTTATGGTCTGGGCTCAACTGGAGAGGAGAAAGATG
ACGTTAGAATACCCCGATATGCACAATGCTGAAATTAGCAGGAGACTGGGCAAGTTATG
GCGGTTACTCGGAGAAGACGAGAAACAACCGTTCATCGAGGAATCCGAGAGGCTACGCA
TCCAACACATGAAGCAATACCCCGACTACAAGTACAAGCCCAAGAGG 3′

B. Translated amino acid sequence:

VKRPMNAFMVWAQLERRKMTLEYPDMHNAEISRRLGKLWRLLGEDEKQPFIEESERLRI
QHMKQYPDYKYKPKR

Figure 3

*Ephydatia muelleri Sox*1

A. Nucleotide Sequence:

5′TCAATCGAGCGCAAGAAGCTAGCCGAGAGAGAACCGCGCTTGCACAACACGCGGTTG
GGCCAGATGTGGAAATGCATGACAGAGGAAGACAAGAAGCCTTTCCGGCTAGAAGCGGA
GAAGCTCAAGACTAAGCTCCTGGAGGAACACCCGG 3′

B1. Translated amino acid sequence:

SIERKKLAEREPRLHNT----RLGQMWKCMTEEDKKPFRLEAEKLKTKLLEEHP

B2. Translated amino acid sequence:

SIERKKLAEREPRLHNTELSKRLGQMWKCMTEEDKKPFRLEAEKLKTKLLEEHP

Figure 4

*Epyhdatia muelleri Sox*2

A. Nucleotide sequence:

5′GTGAAGCGGCCGATGAACGCCTTCATGGTGTGGTCTCGCAAAATGAGAAAGAAGATA
GCGGACGAGAATCCCAAGATGCACAATTCCGAGATCAGCAAACGGTTGGGGGGCGCAGTG
GAAAGCCTTATCGGATGAGGAGAAGCGACCTTACATCGAGGAGGCCAAGCGGCTCAGAG
AGGCGCACATGAAGAAGCATCCAAACTACAAGTACAAGCCCAAGAGGAAGGGC 3′

B. Translated amino acid sequence:

VKRPMNAFMVWSRKMRKKIADENPKMHNSEISKRLGAQWKALSDEEKRPYIEEAKRLRE
AHMKKHPNYKYKPKRKG

Figure 5

*Ephydatia muelleri Sox*3

A. Nucleotide sequence:

5′GCTGGAAAGGAGGAAGATGACCCTGGAATACCCGGACATGCACAACGCGGAGATCAG
CAGGCGCCTGGGCAAGCTGTGGCGGTTACTCACTGATGCAGAAAAGCAGCCCTACGTGG
ACGAGTCGGAGCGGCTAAGAGTGATGCACATGAAGCAGTACCCCG 3′

B. Translated amino acid sequence:

LERRKMTLEYPDMHNAEISRRLGKLWRLLTDAEKQPYVDESERLRVMHMKQYP

C. Nucleotide Sequence:

5′TGGGCAAGCTGGGGCGGTTACTCACTGATGCAGAAAAGCAGCCCTACGTGGACGAGT
CGGAGCGGCTAAGAGTGATGCACATGAAACAGTACCCCGACTACAAGTACAGGCCGCGC
AAGCGCGGGACCAAGAAGACGTGTAAGCAAACCGCGAACAGCAGCGTCGCTGCTGGTGG
TTGCAAACCTTCCACGACGGCGGCAGACTCTCCCTGCGCGCCTTGCGTGTGCGGGAACA
AGACGGCGGAAAAGTGCACGGTGGGTATCCAGTGTACGCTCGATACGAACGGCTCTGAC
ATCATCGAGCGTCACGCTATGGACTGCAGCAGCACCGTCAAGCGTACGGCAGAGATCTC
CATCCAAGTAGGAAATGGTTTGGCCACGGTGAAGGCCGTGACCGCCACAAGTAAACAAC
AGAGCAGCGCGATGCAAACCAAGTACCCGGTCGTCACCGGTGGGAAACGGGTTCGTCTG
GGCTCGGACGGGTCGACAAGACCGCCCAAGCAGGTCAGAAGCGATCAGCAGCACCAGAT
GATGTGCAGATCCCCGACGTTGAGCAAACAGCAGACCAAGGTACCCGACGGTCGCCTAC
CCCTATCTCCGCCGAATTCACTCGACGATCTGGACATGTCTCTCTCGCCAACAGAGGTC
GACCTCTGCCTTCTTCCTGGTCTCAACTTCGTGGACTTGTTGGAACCCATATTCATGTC
GGGGCCTATTCTAACCAGCGCGACCATCTCCCCTCCTCTGAGTGCGACCGGATCAGGAT
TGCAACTGCCGAACTGTTATTCGCCATCGGACATCTCGCAAGATGACAAATCTGTTTTC
GACTTCCCAGACATAAGCCCAGACTTTGCAGAGCTATTCGTCCAAAATCCTTATTCACA
ACTGGACTCGACCATTTCTCCACTTCTCTCGAACTAGTGAGACAGCATACGGTAATTCC
AGCGCCTGTGCTATTCAAATTCAACACCACACAGGATAGGTGACTTGGTGAACACAAAT
ATTCACTTTTTGTGTCACCAAAATAGAAATAAAGTCACTTTCTGTGTATACATGTGTTT
GTACAAACACGTCACTGGATGTGTCCGCGTTTCATCCCGTGCTAGTTGTTCTTGTTCTA
TGATCACTTTATGCATCATTTGAATTGCCACAAAAAAAAGAAAAAGACATAACGAAGAT
CTATTGCTCCCACGCGTTGG 3′

D. Translated amino acid sequence:

GKLWRLLTDAEKQPYVDESERLRVMHMKQYPDYKYRPRKRGTKKTCKQTANSSVAAGGC
KPSTTAADSPCAPCVCGNKTAEKCTVGIQCTLDTNGSDIIERHAMDCSSTVKRTAEISI
QVGNGLATVKAVTATSKQQSSAMQTKYPVVTGGKRVRLGSDGSTRPPKQVRSDQQHQMM
CRSPTLSKQQTKVPDGRLPLSPPNSLDDLDMSLSPTEVDLCLLPGLNFVDLLEPIFMSG
PILTSATISPPLSATGSGLQLPNCYSPSDISQDDKSVFDFPDISPDFAELFVQNPYSQL
DSTISPLLSN**DSIR*FQRLCYSNSTPHRIGDLVNTNIHFLCHQNRNKVTFCVYMCLY
KHVTGCVRVSSRASCSCSMITLCII*IATKKRKRHNEDLLLPRV

Figure 6

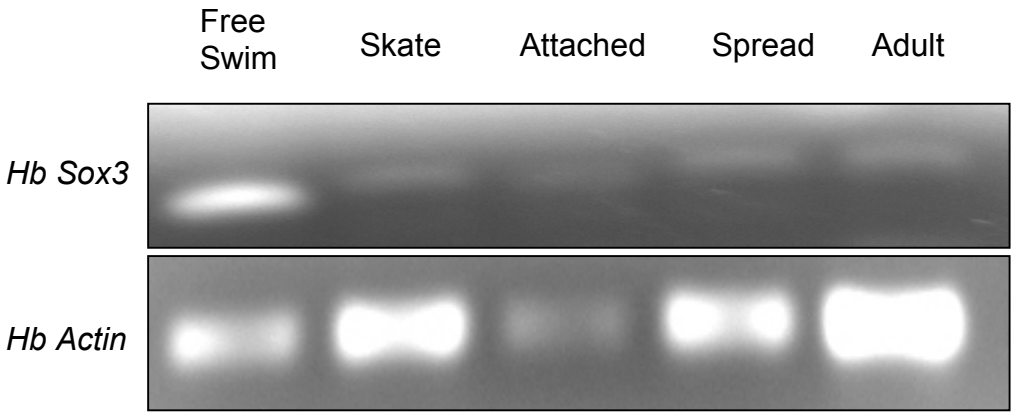

Figure 7

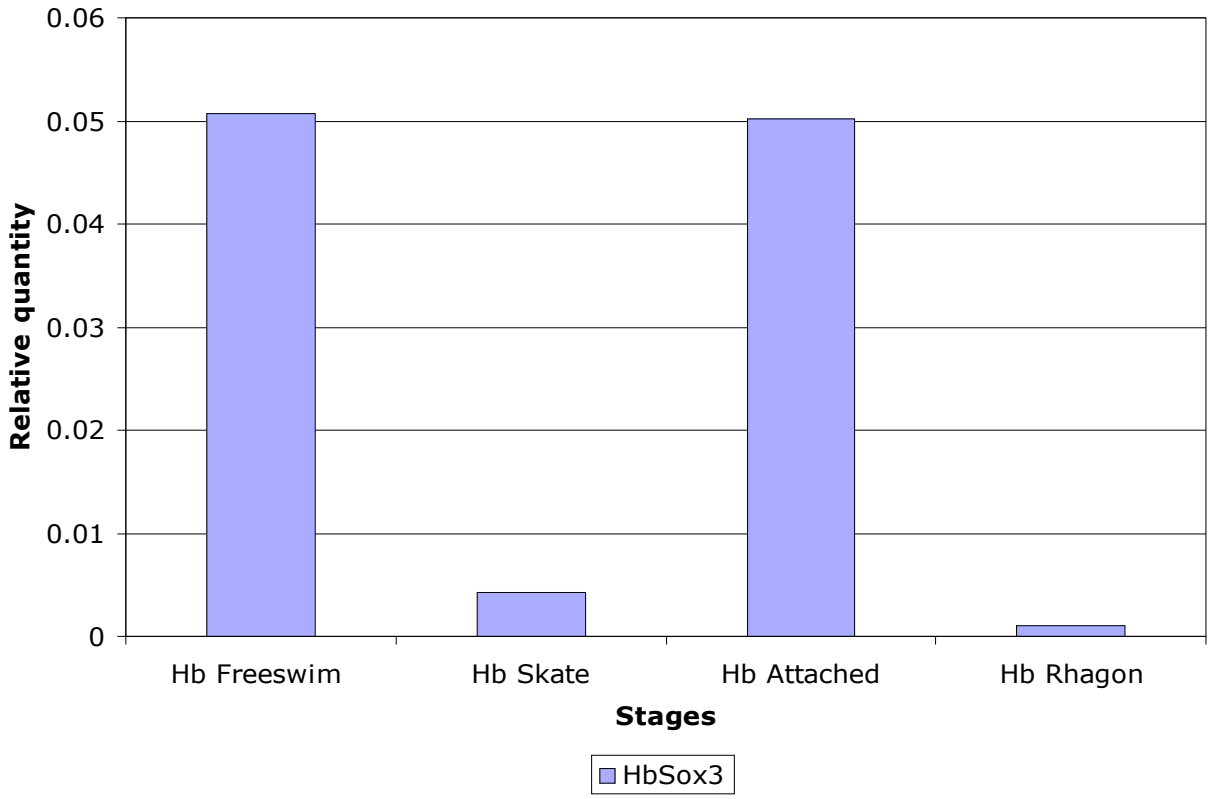

Figure 8

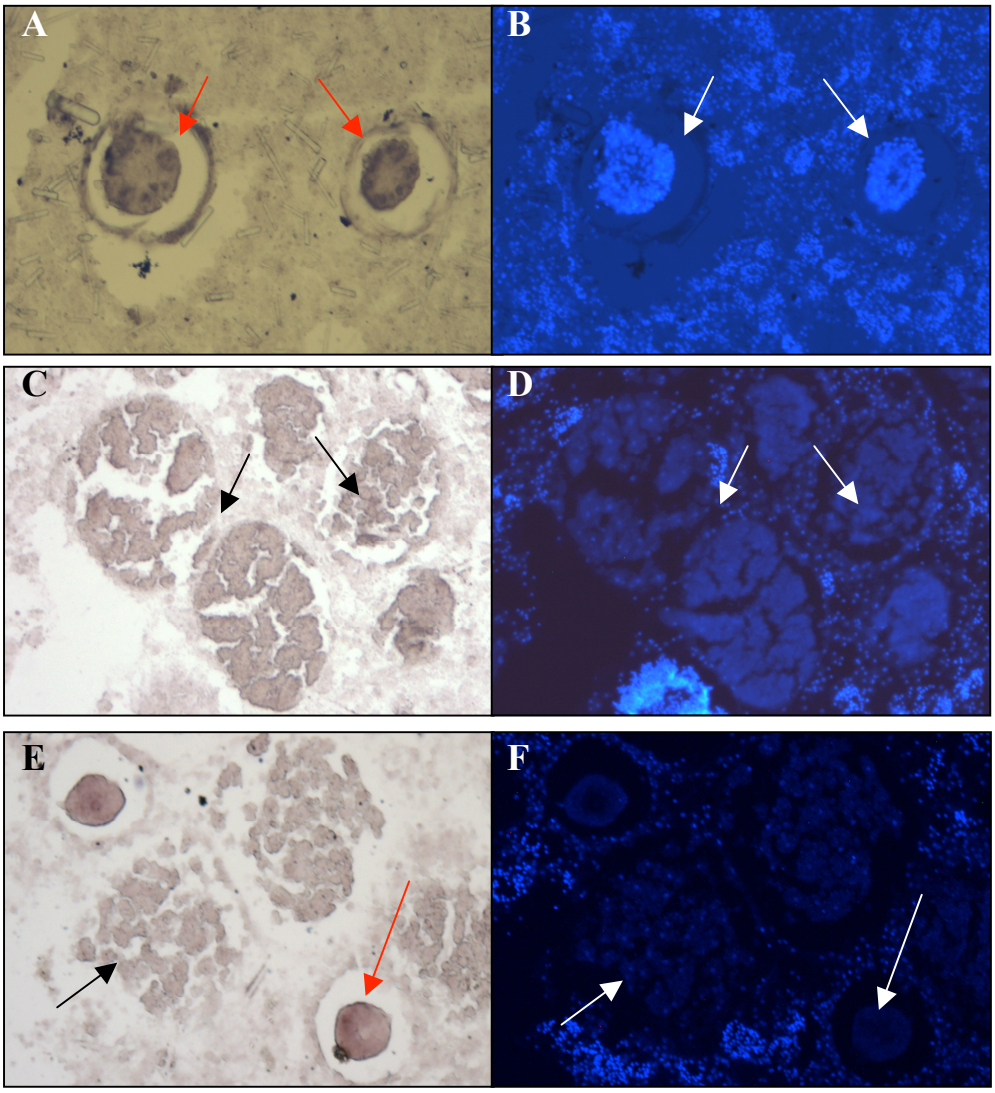

Figure 9

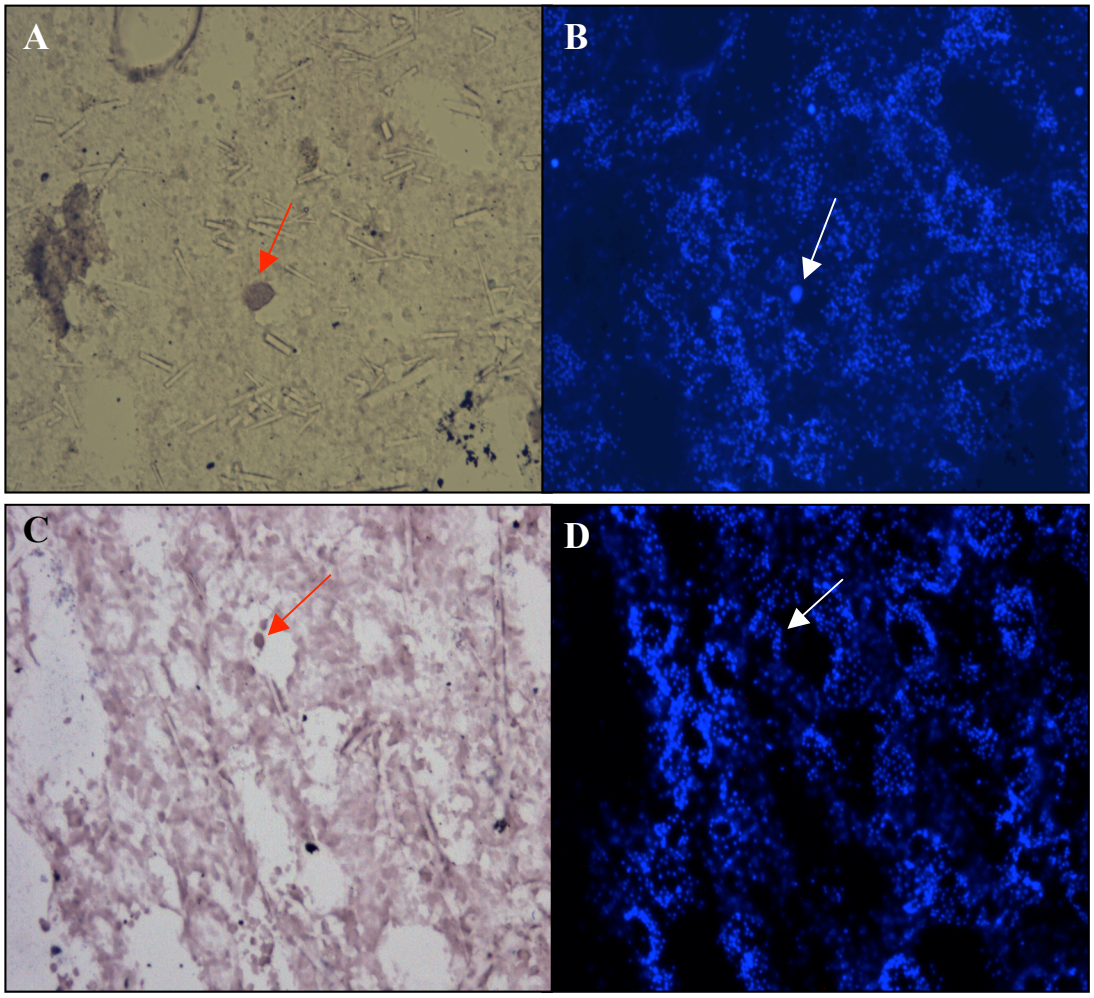

Figure 10

Figure 11

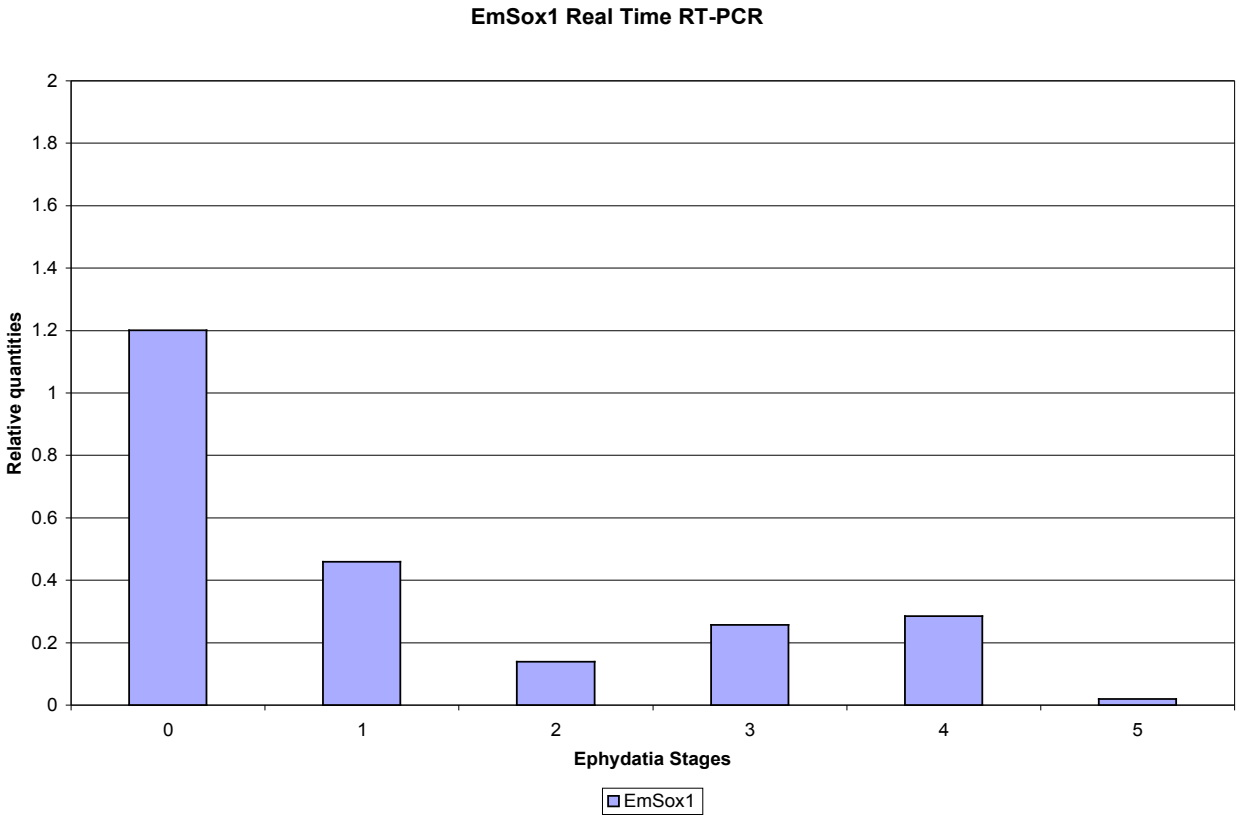

Figure 12

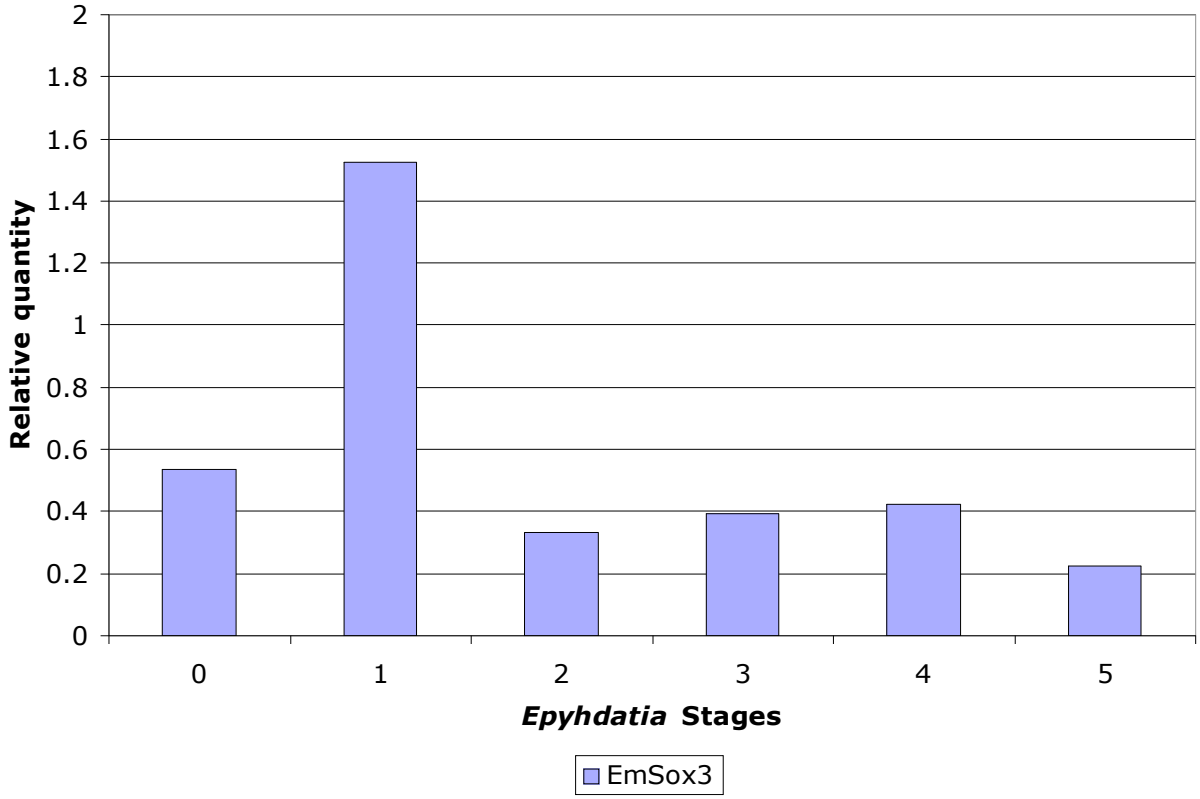

Figure 13

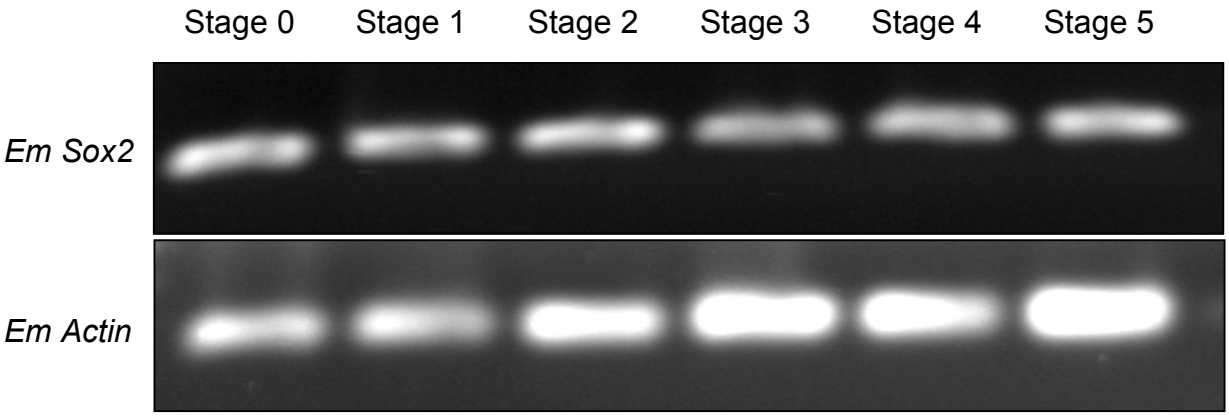

Figure 14

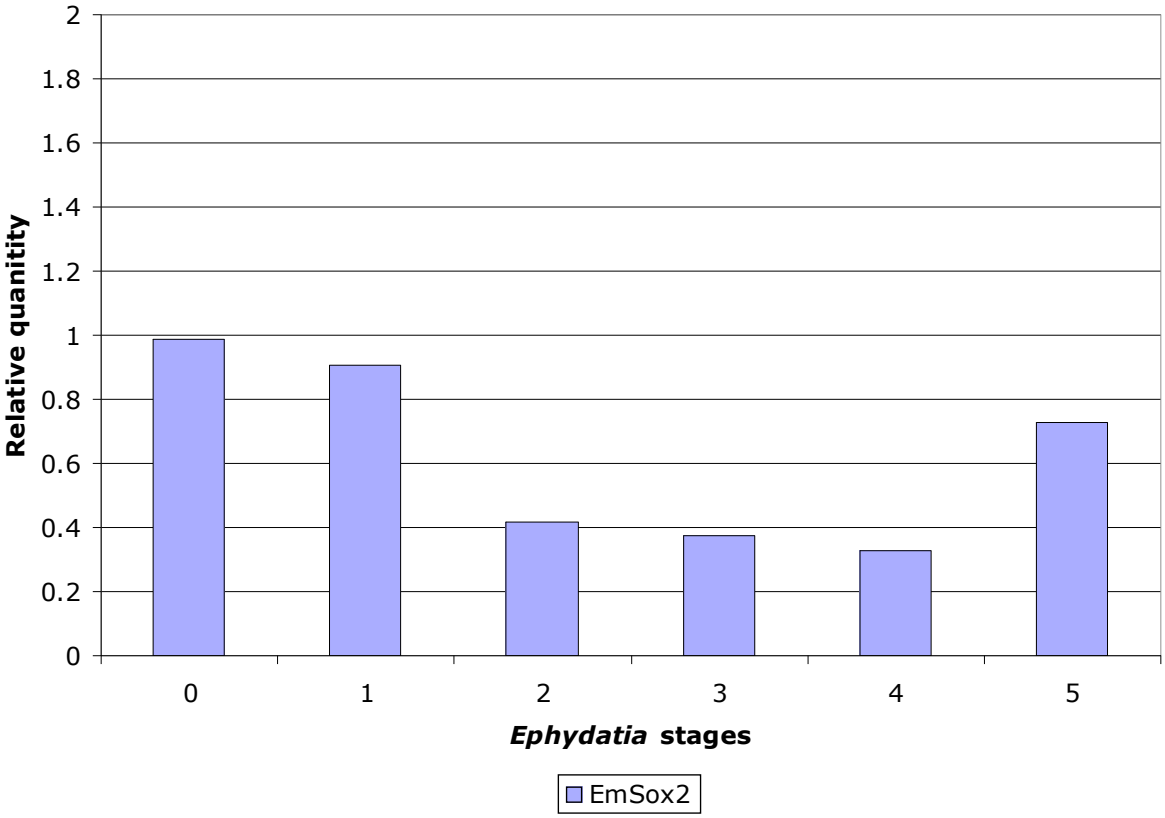
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Relative quanitity
0
1
2
3
4
5
Ephydatia stages
EmSox2

Figure 15