5-2015

# Cloning of the SOX2 gene from Diplosoma listerianum

Jessica M. Simonelli
jsimonelli@my.uri.edu

Steven Q. Irvine
*University of Rhode Island*, sirvine@uri.edu

# Cloning of the SOX2 gene from *Diplosoma listerianum*

Jessica Simonelli[1], Dr. Steven Q. Irvine[2]

1Coastal Fellow, The University of Rhode Island 2Department of Biological Sciences University of Rhode Island

## Background

### Asexual Development in the tunicate species *Diplosoma listerianum*

Tunicates, more commonly known as "sea squirts" or ascidians, are filter feeding marine invertebrates. *D. listerianum* is a colonial ascidian, meaning that thousands of individual animals live together in a colony, surrounded by a tough outer matrix, and can reproduce sexually as well asexually through a process known as pyloric budding. In this process, the esophagus of a single animal will form two buds that can each develop into a whole new animal, allowing for very rapid reproduction. Pyloric budding is a unique feature of a few tunicate species including *D. listerianum*. The overall goal of my research is to learn more about the genes that encode this organism's asexual development.

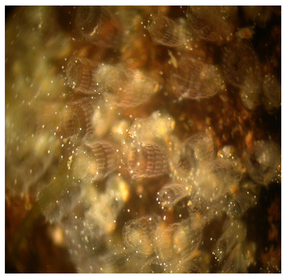

**Figure 1: close up of D. listerianum**


**Figure 2: Buds of D. listerianum- the arrow is pointing to the pharyngeal bud**

### Significance of the SOX2 transcription Factor

SOX2 is an important transcription factor (transcription factors help promote the expression of genes) in embryonic stem cells. SOX2 is present in many organisms (including humans) and is highly conserved, meaning the sequence is very similar in different species. It is essential for pluripotency in stem cells, which is the ability of a stem cell to differentiate into different tissue types. From my own past research, I know that the SOX2 gene is expressed in *Diplosoma*, but the role that it plays in this organism has not been previously studied. Our hypothesis is that this gene is important for *Diplosoma*'s asexual development.

### In Situ Hybridization Determines Where A Gene is Expressed in a Fixed Tissue Sample

In order to determine the role SOX2 plays in *Diplosoma*, we must look at where gene expression occurs in the organism. If the gene is important for asexual development, it will most likely be expressed in the budding tissue. In order to determine where gene expression occurs, we can use a method known as "in situ hybridization"

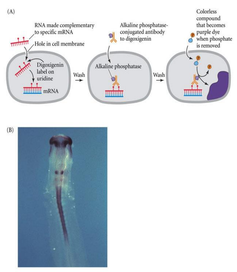


**Figure 3: Whole Mount In Situ Hybridization in a Fixed Chick Embryo. The antisense probe hybridizes with the mRNA in the embryo, causing a dark precipitate to form where gene expression occurs (Source: http://10e.devbio.com/)**

## Introduction

*Diplosoma listerianum* is an invasive tunicate species that can reproduce asexually by pyloric budding. Our research primarily focuses on the genetic basis for the asexual development in this organism. From past research I have found that *Diplosoma listerianum* expresses a gene that codes for the SOX2 transcription factor, and successfully sequenced a portion of this gene. The end goal of this project is to determine where gene expression occurs in the organism. If we find that gene expression occurs in the budding tissue, it will provide evidence for the fact that this gene is involved in asexual development in *Diplosoma*. One way we can do this is to use a method known as in situ hybridization to localize where gene expression occurs. This project accomplished the beginning steps of this process, allowing for future research to determine where SOX2 gene expression occurs in *Diplosoma*.

## Methods

### Overall Goal

The goal of in situ hybridization is to find where in a fixed tissue sample a gene is being expressed. This is done by creating a labeled complementary RNA "probe" that will hybridize or basepair to a specific mRNA sequence within the organism. If our hypothesis is correct, our probe will hybridize to an mRNA sequence in the budding tissue of *Diplosoma.* The first step in this process was to insert the SOX2 gene in a plasmid vector, in order to definitively determine the sequence of the fragment.

### Results from Past Research

My past research involved testing to see which developmental genes are expressed in *Diplosoma*. In order to do this, I extracted messenger RNA from mature *Diplosoma* zooids and made complementary DNA. Then, I used several different primers (a primer is a sequence that initiates DNA synthesis) to create copies of specific genes that code for development in other species, to see if these genes are also present in *D. listerianum*. Using a method known as "polymerase chain reaction," (PCR), I used these primers to make many copies of the genes, and had them sequenced to determine the order of the nucleotides that make up the gene. From doing this, I found that the gene encoding the SOX2 transcription factor is expressed in *Diplosoma*, and determined the sequence of part of the gene. I then searched for the sequence using the program BLAST, which could compare my sequence to the ESTs (expressed sequence tags) of previously sequenced genomes of other organisms. The program indicated that the sequence from our PCR product is most closely similar to genes in the SOX family. Our sequence is highly similar to the sequence that codes for SOXB1 in *Ciona intestinalis*, another tunicate species.

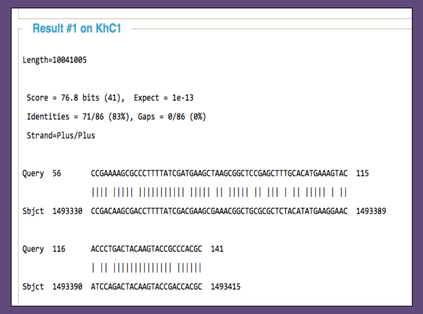


**Figure 4: Comparison of the *Diplosoma* SOX2 sequence to that of *Ciona intestinalis*. An 83 % match was found between the two sequences. (Source: ANISEED Database)**

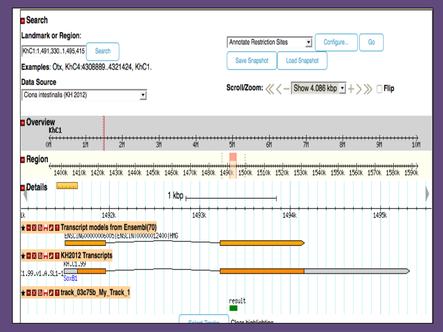


**Figure 5: Alignment of SOXB1 sequence in *Diplosoma* compared to its location in *Ciona Intestinalis* (Source: ANISEED Database)**

### Cloning the SOX2 gene for the in situ hybridization

Since the fragment of the SOX2 gene I managed to sequence was only 200 basepairs in length, the first step in this project is to sequence a large enough portion of the gene to create the probe for the in situ hybridization (this method typically requires the probe to be at least 600 basepairs in length). In order to do this, I needed to get a more exact sequence for the fragment I managed to clone. Therefore, I inserted the gene of interest into a "vector", which is a DNA molecule that can be transferred into cells and replicated so it can be cloned and sequenced. The vector I used is an artificially constructed plasmid known as puc18. A plasmid is a circular DNA molecule commonly found in bacterial cells that is able to replicate itself. Therefore, plasmids are often used as cloning vectors for specific sequences that can be inserted in the plasmid at a site called the "multiple cloning site." A copy of the SOX2 gene was inserted within the plasmid DNA and sealed using an enzyme called DNA ligase. This generated what we refer to as a "recombinant plasmid," which now contains the gene of interest (in this case SOX2) along with the rest of the plasmid DNA. We can then use this to make our probe, which is the "antisense" RNA sequence. In a process known as "transformation," we transferred the recombinant plasmid DNA into *E. coli* cells, which we grew on plates containing the antibiotic ampicillin. Since the puc18 plasmid contains a gene that codes for ampicillin resistance, the cells that have taken up the plasmid DNA will survive while the rest will be killed off. That way, we were able to only grow colonies that contained our vector. We then did a plasmid preparation to extract and purify the plasmid DNA from the cell, and sequenced it. I can later use this sequence to design a primer that will allow me to clone and sequence a larger fragment of the gene.

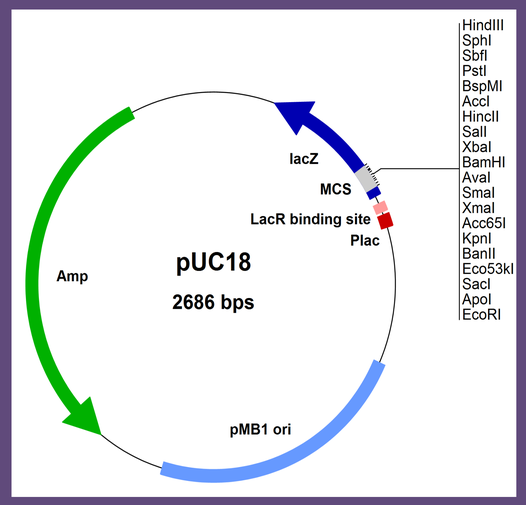


**Figure 6: Map of the puc18 plasmid. The MCS stands for "multiple cloning site and the gene labelled "amp" is the ampicillin resistance gene (Source: http://www.mobitec.com)**

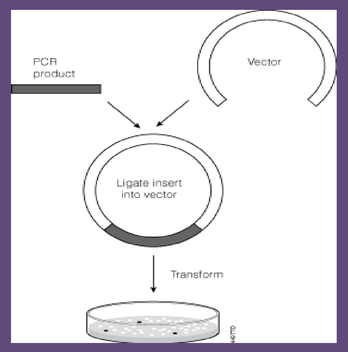


**Figure 7: Ligation of the PCR product into the plasmid vector (source:https://promega.wordpress.com)**

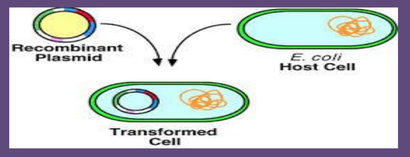


**Figure 8: Transformation of the recombinant plasmid into *E. Coli* cell (Source: http://www.biology.arizona.edu/)**

## Data and Results

### Results from Sequencing

I sequenced the recombinant plasmid DNA from six colonies grown on two different ampicillin plates- three of the colonies were grown on one plate, and the other three were grown on the other. I obtained two distinct sequences for the colonies grown on each of the separate plates- three of the nucleotides were different in the two sequences. When I searched the sequences using the program BLAST, both were shown to be most closely similar to SOX genes, though the results had definite differences. Therefore, it is possible that I actually obtained the sequence of two genes in the SOX family. There are many different SOX genes that code for transcription factors essential for development in many organisms. These genes are closely related but have different functions. Thus, I may be able to create two different primers, and determine a larger fragment of the sequence of two different genes within the SOX family

**Sequence 1 Top Blast Results**

**Nucleotide Sequences Comparison:**

| Description | E Value | Identity | Accession Number |
|---|---|---|---|
| Cyprinus carpio genome assembly common carp genome, scaffold LG35 | 4e-34 | 77% | LN590686.1 |
| Xenopus laevis SRY (sex determining region Y)-box 21 (sox21), mRNA | 4e-34 | 77% | NM_001172213.1 |
| PREDICTED: Micropilitis demolitor transcription factor Sox-2 (LOC103569292), mRNA | 5e-33 | 76% | XM_008546519.1 |
| PREDICTED: Callorhinchus milii SRY (sex determining region Y)-box 2 (sox2), mRNA | 5e-33 | 76% | XM_007907767.1 |
| PREDICTED: Sarcophilus harrisii SRY (sex determining region Y)-box 21 (SOX21), mRNA | 5e-33 | 76% | XM_003765793.1 |

**Comparison to Proteins in *Homo sapiens* and *Ciona intestinalis***

| Description | E Value | Identity | Accession Number |
|---|---|---|---|
| Transcription Factor SOX-14 [Homo sapiens] | 3e-36 | 87% | NP_004180.1 |
| Transcription Factor Sox2 [Homo sapiens] | 6e-36 | 85% | 2LE4_A |
| HMG transcription factor SoxB2 [Ciona intestinalis] | 2e-35 | 85% | NP_001122329.1 |
| SoxB1 protein [Ciona intestinalis] | 7e-34 | 81% | CAD58840.1 |

**Sequence 2 Top Blast Results:**

**Nucleotide Sequences Comparison:**

| Description | E Value | Identity | Accession Number |
|---|---|---|---|
| PREDICTED: Sarcophilus harrisii SRY (sex determining region Y)-box 21 (SOX21), mRNA | 5e-33 | 76% | XM_003765793.1 |
| Xenopus laevis SRY (sex determining region Y)-box 21 (sox21), mRNA | 4e-34 | 76% | NM_001172213.1 |
| PREDICTED: Tribolium castaneum transcription factor Sox-3 (LOC663351), mRNA | 2e-32 | 77% | XM_008194925.1 |
| PREDICTED: Pogonomyrmex barbatus transcription factor Sox-19b (LOC105434020), mRNA | 2e-31 | 76% | XM_011649574.1 |
| Cyprinus carpio genome assembly common carp genome, scaffold LG35 | 2e-31 | 76% | LN590686.1 |

**Comparison to Proteins in *Homo sapiens* and *Ciona intestinalis***

| Description | E Value | Identity | Accession Number |
|---|---|---|---|
| Transcription Factor SOX-14 [Homo sapiens] | 2e-35 | 85% | NP_004180.1 |
| Transcription Factor Sox2 [Homo sapiens] | 4e-35 | 84% | 2LE4_A |
| HMG transcription factor SoxB2 [Ciona intestinalis] | 1e-34 | 84% | NP_001122329.1 |
| SoxB1 protein [Ciona intestinalis] | 6e-33 | 79% | CAD58840.1 |

**Figure 9: The figure above shows the top five blast results for each of the nucleotide sequences , and compares the translated proteins to the closest results in humans and another tunicate species, *Ciona intestinalis*. (Source: ANISEED database)**

### The Next Steps:

The next step in this project is to carry out the rest of the in situ experiment by creating the antisense RNA probe that will hybridize to the mRNA in *Diplosoma* adults that have been fixed in formalin.

**Explanation:**

When DNA is translated to make a protein, it must first be transcribed into what is referred as "messenger" RNA or mRNA. This messenger RNA is then processed and taken to a cell organelle known as a ribosome, where it provides the information for the order of amino acids that make up a protein. The mRNA for a particular gene will only be present in cells where that gene is being expressed. Thus, to see where a particular gene is being expressed in an organism we can make a probe that will hybridize to that specific mRNA sequence. This probe must be complementary to the mRNA,. The probe can then be labeled and added to our fixed sample of Diplosoma adults, to see where it hybridizes. If it hybridizes to mRNA in the budding tissue, this may be evidence that SOX2 may play a role in the asexual development of this organism.

**Future research:**

Future research can focus on whether other genes that are essential for stem cell pluripotency are also important for asexual reproduction in *Diplosoma*. Examples of such genes would include other SOX genes, Oct genes, and VASA genes.

## Acknowledgements

This project was supported by a URI Council for Research Proposal Development Grant.
I would also like to thank the Coastal Fellows program at the University of Rhode Island