universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Assessing the role of Myc in the marine sponge *Suberites domuncula*: Towards the establishment of a technique for gene function analysis“

verfasst von / submitted by

Maximilian Traun, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2021/ Vienna 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 877
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Genetik und Entwicklungsbiologie
Betreut von / Supervisor:	Dr. Florian Raible
Mitbetreut von / Co-Supervisor:	Roger Revilla-i-Domingo, PhD

Abstract

As one of the earliest branching animal phyla, sponges play a very important role in discovering ancestral metazoan molecular mechanisms. In particular, they are of interest with regards to questions about the origin of multicellularity as well as the evolution of stem cells. One interesting hypothesis claims that the transcription factor Myc played an important role in regulating stem cell functions in early animals. If confirmed, this would have important implications for our understanding of the evolution of multicellularity.

The current model for the sponge stem cell system claims that one specific sponge cell type, the archaeocytes, act as totipotent stem cells, suggesting that they have the capacity to self-renew and to differentiate into all cell types of the animal. Interestingly, these archaeocytes are enriched for the expression of Myc.

In *Drosophila* and mammals, Myc has been shown to regulate important stem cell functions, such as proliferation, differentiation and apoptosis, and a recent study reported 100 *bona fide* genes directly regulated by Myc in a human cell line. Many of these genes can be directly linked to well-studied stem cell functions of Myc. The availability of such direct target genes provides a set of genes that can be used for assessing a potential ancient role of Myc factors in stem cell regulation.

Using the demosponge *Suberites domuncula* as a model organism, in this thesis, I have analyzed the expression of the sponge homologs of a subset of the genes shown to be directly regulated by Myc in human, and found that about half of the analyzed genes are significantly co-expressed with *myc* RNA in the sponge. This suggests a conserved regulatory network, strengthening the hypothesis about an ancestral role of Myc in early animal stem cells. In addition, I have performed work that aimed to establish a reliable technique to modulate Myc RNA and/or protein levels and thereby test the function of Myc during sponge regeneration. I tackled this challenge by trying to establish four techniques to modulate gene expression, with the intention of analyzing the effect of elevated or decreased Myc expression. These four techniques are commonly used in other model organisms: antisense morpholino knockdown, RNAi, Myc overexpression and dominant negative Myc expression. Even though I was not able to establish any of them as an efficient way to knock down genes in my model organism, this work provides a set of reagents and established assays that may be useful for the establishment of a technique to modulate gene expression in *Suberites domuncula* in the future.

Zusammenfassung

Schwämme spielen als Vertreter eines der basalsten Tier-Phyla eine wichtige Rolle bei der Entdeckung ursprünglicher molekularer Mechanismen. Besonders bei Fragestellungen über die Entstehung der Mehrzelligkeit und der Entwicklung von Stammzellen sind sie von besonderem Interesse. Eine interessante Hypothese besagt, dass der Transkriptionsfaktor Myc eine wichtige Rolle bei der Kontrolle von Stammzellmechanismen in frühen Tieren gespielt hat. Sollte sich das bewahrheiten, hätte das wichtige Implikationen im Hinblick auf unser Verständnis von Mehrzelligkeit.

Das gängige Schwamm-Stammzellmodel besagt, dass ein bestimmter Zelltyp in Schwämmen, die Archaeocyten, als totipotente Stammzellen fungieren, was bedeutet, dass sie sich selbst erneuern können und zu jedem anderen Zelltyp im Schwamm werden können. Interessanter Weise ist Myc besonders in diesen Zellen angereichert.

In *Drosophila* und Säugetieren wurde bereits gezeigt, dass Myc so wichtige Stammzellfunktionen erfüllt wie Proliferation, Differenzierung und Apoptose, und eine vor kurzem veröffentlichte Studie berichtet von 100 Genen, welche in einer menschlichen Zelllinie direkt von Myc kontrolliert werden. Interessanter Weise werden viele dieser Gene direkt mit gut untersuchten Stammzellfunktionen von Myc assoziiert. Die Verfügbarkeit solcher direkter Zielgene kann zur gezielten Untersuchung einer potentiell ursprünglichen Rolle von Myc in Stammzellmechanismen genutzt werden.

In dieser Masterarbeit habe ich im Schwamm *Suberites domuncula* die Expression von Homologen jener Myc-regulierten Gene analysiert, mit dem Befund, dass ungefähr die Hälfte der analysierten Gene signifikant mit *myc*-RNA zusammen exprimiert werden. Dies legt ein gut konserviertes regulatorisches Netzwerk nahe, und stärkt damit die Hypothese, dass Myc eine ursprüngliche Rolle in den Stammzellen früher Tiere hat. Außerdem habe ich darauf hingearbeitet, eine verlässliche Methode zur Modulation vom Expressionsniveau von *myc* in Schwämmen zu etablieren und damit die Rolle von Myc bei der Regeneration von Schwämmen zu untersuchen. Dafür versuchte ich vier verschiedene Techniken zur Modulierung von Genexpressionsniveaus zu etablieren, die regelmäßig in anderen Modellorganismen genutzt werden: antisense vivo-morpholino knockdown, RNAi, Überexpression und Expression eines dominant negativen Myc-Konstrukts. Obwohl ich letztlich nicht in der Lage war, reproduzierbare Änderungen der Genexpression in *Suberites domuncula* zu erreichen, habe ich doch eine Reihe an Reagenzien und Prüfverfahren etabliert, welche künftig bei der Etablierung einer Modulation des Genexpressionsniveau in *Suberites domuncula* von Nutzen sein könnten.

Acknowledgments

This thesis would not have been possible if not for a lot of people whom I would like to thank here:

First, I would like to thank Dr. Florian Raible for providing me with the space and funding to work at the Max Perutz Laboratories as well as offering advice and support when needed. I have learned much while in his work group and will continue to benefit from the very positive environment fostered under his guidance.

Especially thankful I am for my co-supervisor Dr. Roger Revilla-i-domingo who quickly trusted me to work on my own while always being available to help and offer advice. During frustrating times, he motivated me to keep positive, when I made mistakes he helped rather than just correcting me and made me feel treated fairly overall. For this I am grateful for I was able to do my best work without feeling pushed over the limit, even during difficult times.

I also want to thank all the other lab members, for the enjoyable time and the help and advice I sometimes greatly needed.

Content

1. Introduction

Sponges (Porifera) are one of the earliest branching metazoan phyla (**Fig. 1**).[1–3]. They are filter feeding organisms and have a very simple morphology. They consist of only approximately 10 cell types[4], and lack muscles, organs and a nervous system.[5] In spite of their simplicity, they have a remarkable regeneration capacity. This makes them ideal to study ancestral metazoan features that are relevant for our understanding of the origin of multicellularity as well as the origin of stem cells.

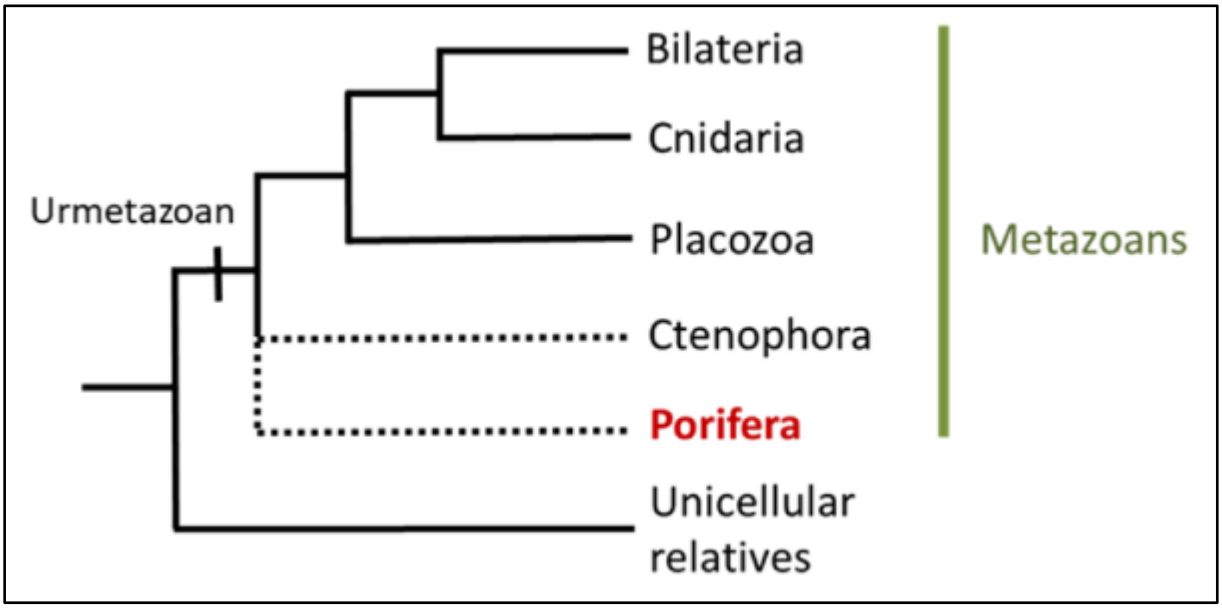


Figure 1: Phylogenetic tree of early metazoa.

Sponges represent either a sister group to all animals or all animals minus the Ctenophora. (C Schmidt, 2018 "Establishment of transgenesis in the demosponge Suberites domuncula")

1.1 Sponge cell types and general structure

The body plan of a sponge consists of three parts[4]: the outer epithelium, the inner epithelium and the mesohyl. The outer epithelium is made up of exopinacocytes and basopinacocytes while the inner epithelium consists of mostly endopinacocytes with chambers for the flagellated feeding cells, the choanocytes. These choanocytes produce a water flow with their flagellum and filter nutrients out of the water. The other six morphologically well characterized cell types live in the mesohyl. One of these will be the most prominent cell type in this thesis: the archaeocytes.

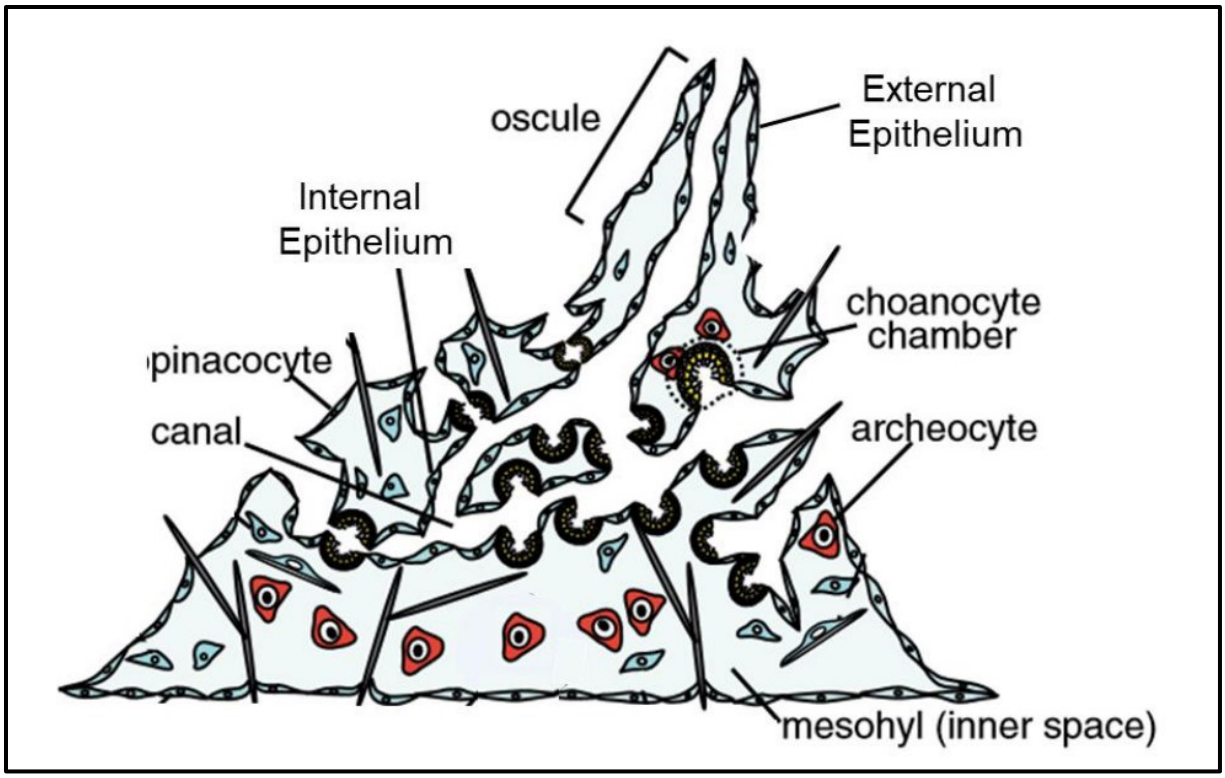


Figure 2: Basic buildup of a sponge with relevant cell types. (Adapted from N. Funayama, 2013)

Our model organism, *Suberites domuncula,* is a marine demosponge suitable for growth and regeneration experiments[6] and can easily be kept in an enclosed seawater aquarium. It has been shown before that pieces cut out of the sponge can be cultured for extended periods of time and regenerate body structures like canals and a new oscule[6], *i.e.* the opening through which the water circulating within the animal during feeding, flows out of the animal.

1.2 Regeneration and stem cells in sponges

Sponges have strong regeneration capabilities[4,7], and multiple findings support the hypothesis that archaeocytes function as stem cells and are the basis for regeneration in sponges:

- Cells from dissociated sponge tissue can aggregate and form a new small sponge[8]. This effect gets lost when archaeocytes are eliminated from the dissociation.
- According to one study, regeneration after surgical removal of a portion of apical ectosome (superficial part of the sponge), along with a directly underlying section of the aquiferous system (choanocyte chambers and canals), involved predominantly archaeocytes as well as transdifferentiated choanocytes[9].
- Sponges can reproduce sexually or asexually by producing gemmules [10]. These gemmules consist of a mass of archaeocyte-like cells and are able to grow into an adult sponge again.

- Archaeocytes have been suggested to have totipotent stem cell properties[4,11–14], meaning that they are able to self-renew and to differentiate into all sponge cell types.(**Fig. 3**). Although the totipotency of archeocytes has not been experimentally demonstrated so far, single-cell mRNA-sequencing experiments suggest differentiation trajectories between the archeocytes and many of the differentiation cell types[15].

Furthermore, in more recent studies markers like *vasa*[16] and *piwi*[12] were found to be highly enriched in archaeocytes of several sponge species[6,15–18]. *piwi and vasa* are genes active in vertebrate[19,20] germ cells as well as basal metazoan[21–23] somatic cells that give rise to the germline and they are part of the germ-line multipotency program.

Although still controversial, some studies have suggested that choanocytes also exert some stem cell functions: they are proliferative, they can give rise to gametes and it has been suggested that they can dedifferentiate into archeocytes[11].

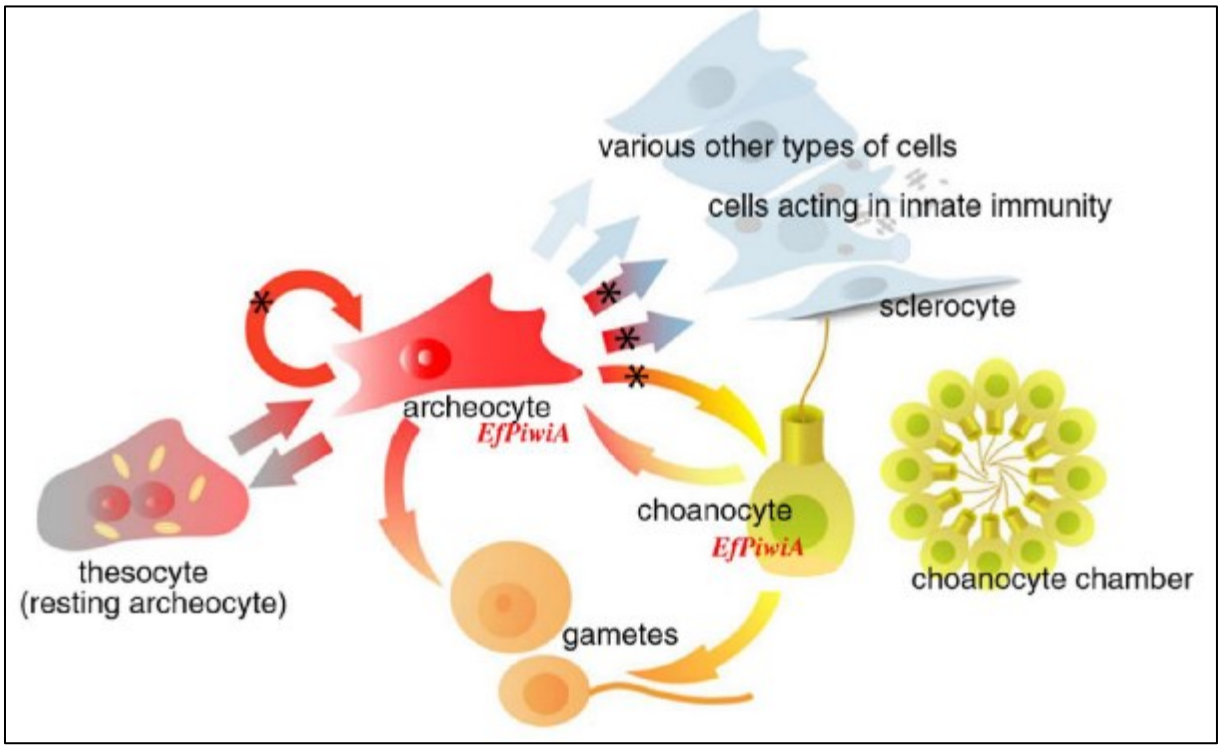


Figure 3: The current view of the stem cell system in sponges by N. Funayama, 2013.

Archaeocytes are suggested to be totipotent stem cells, *i.e.* that give rise to other cell types. As stem cells they can self-renew and also produce female gametes while spermatozoa originate from choanocytes. Choanocytes have been suggested to be able to dedifferentiate into archaeocytes under specific conditions.

1.3. Myc is an important regulator of stem cell mechanisms

Stem cells are cells in multicellular organisms able to self-renew and differentiate into other cell types. Stem cells also often rely on glycolysis for energy production because of their

localization in hypoxic environments[24]. Myc proteins are part of a protein family that, in mice and *Drosophila* have been shown to be involved in the regulation of a plethora of stem cell mechanisms like proliferation[25], stem cell differentiation[26] and glycolysis[27].

The most prominent *myc* genes in mice and other vertebrates are *c-myc, N-myc* and *L-myc* with c-*myc* being the most abundant and best studied. It has also been shown that N-Myc can substitute c-Myc function, suggesting at least partial redundancy between genes within this family. Furthermore, the expression of the *c-myc* gene is tightly controlled, and its deregulation and amplification is frequent in mammalian cancers. In a previous transcriptomic analysis in *Suberites* in our laboratory, *myc* was found to be enriched in archaeocytes, alongside other typical stem cell genes like *vasa, pl-10*, *piwia and piwib.*

c-Myc in vertebrates typically consists of three major domains: a trans activation domain (TAD), a basic helix-loop-helix (bHLH) domain as well as a leucin zipper (LZ) domain.[28] The TAD is located at the N-terminus of the protein and contains regions which are responsible for transcription activation functions. The bHLH as well as the LZ domain are situated at the C-terminal end of the protein and facilitate DNA binding and hetero-dimerization with Mycs major binding partner Max[29].

Some studies have suggested that Myc works as a general amplifier of gene expression[30,31]. Contrary to this model, a recent study has reported 100 specific direct transcriptional targets of Myc, by developing a new technique SLAM-Seq in a human cell line[32], monitoring all the RNA being synthesized after a certain point. Combining this with targeted protein depletion using auxin inducible degradation to eliminate existing Myc gives strong evidence of direct transcriptional control of Myc for these genes. This strongly suggests that Myc has a wide range of specific targets opposed to the model that Myc generally enhances transcription and translation[33]. Importantly, many of the 100 reported direct targets of Myc can be directly linked to the well-studied functions of Myc in stem cells[32]. For example, Nme1 and Nme2 are involved in nucleotide synthesis[34] needed for RNA and DNA synthesis, Tfap4 regulates genes controlling cell proliferation[34] and Rpl14[34], Nifk[34] and Nolc1[34] are involved in ribosome biogenesis.

Interestingly, Myc has been shown to be enriched in the archaeocytes of several sponge species[6,15–18], and it has been suggested to control proliferation in a unicellular organism closely related to animals[35]. These observations led to the hypothesis that some of the functions of Myc in *Drosophila* and mammal stem cells might be ancestral, and that Myc might have played a key role in the evolution of stem cells and multicellularity[18].

1.4. Goals of this thesis:

Although we now know the genes expressed in most sponge cell types, to our knowledge no efforts have been reported so far aiming at establishing candidate direct targets of Myc in sponges. In addition, techniques to analyze the molecular function of genes are still very limited[6] in sponges, making it impossible to test the hypothesis that the function of Myc is conserved between sponges and other animals.

To address these limitations, I have followed two strategies:

1. I exploited the availability of single-cell mRNA-sequencing data in the marine sponge *Suberites domuncula* to test whether the sponge homologs of the direct targets of Myc reported in Muhar et al. are co-expressed with Myc in *S. domuncula*. This strategy also involved bioinformatic analyses to validate the identify of S. domuncula *myc* and *max* genes as a proper orthologs of their vertebrate counterparts.

2. I have worked towards the establishment of a reliable gene knockdown or overexpression technique in *S. domuncula,* with the aim to investigate the role of Myc in sponge stem cells. This thesis has sought to establish the following techniques of gene knockdown: antisense vivo-morpholino and RNAi using small hairpin RNAs (shRNAs), as well as to otherwise modulate the levels of Myc by overexpression, or to modulate its function by expressing a dominant negative form of Myc.

2. Materials and methods

2.1 Animals

Specimens of the marine sponge *Suberites domuncula* (**Fig. 4**) were cultivated at 15 – 16 °C in an enclosed seawater aquarium.

Figure 4: Model organism: *Suberites domuncula*.

The demonsponge *Suberites domuncula* grows on snail shells that are inhabited by hermit crabs.

2.2 Identification of *S. domuncula myc* and *max*

To identify candidates of *S. domuncula myc* and *max* genes, I first collected the available Myc protein sequences from several animals (including two sponge species): *Homo sapiens, Mus musculus, Gallus gallus, Oryzias latipes, Hydra vilgaris, Amphimedon queenslandica, Ephydatia fluviatilis.* I then used a similarity search (tblastn) against the available *S. domuncula* transcriptome[6], and selected all *S. domuncula* transcripts with a tblastn E value < 1e-4. A phylogenetic tree was then constructed using IQ TREE. To do this, first the derived protein sequences were aligned using MUSCLE after which the output file containing the aligned sequences, was converted into a .phy file using CLUSTALW. This file could then be used to run IQ TREE. The general string to run IQ TREE looked as follows: iqtree -s Filename.phy -m TEST -alrt 1000 -bb 1000. The bootstrapped output file with the suffix .phy.contree was then uploaded to iTOL for tree visualization. As an outgroup for the phylogenetic tree I used the gene identified to be most similar to Myc in *Mus musculus*, namely, the transcription factor Max. The *S. domuncula max* gene was obtained similarly.

2.3 Characterization of *S. domuncula* Myc and Max protein domains

For analysis of the domains of the collected Myc and Max protein sequences I used the protein domain prediction program Pfam[36].

2.4 Identification of *S. domuncula* homologs of mouse genes

Candidate *S. domuncula* homologs of mouse proteins were obtained by comparing the available transcriptome of *Suberites domuncula*[6] to the mouse proteins using BLAST. The best hit (with an E value threshold of 1e-4) was then checked reciprocally by BLASTing the corresponding S. domuncula protein sequence against the human proteome. The candidate S. domuncula transcripts were considered a bona fine homolog only if the expected mouse protein was recovered as the best hit in our reciprocal Blast.

2.5 Gene co-expression analysis

Expression of *myc* and genes of interest in *S. domuncula* single cells was obtained from available 10xGenomics single-cell mRNA sequencing data. We first counted the number of cells that show overlapping expression between *myc* and each gene of interest. To calculate the probability (p value) that the number of cells co-expressing *myc* and each gene of interest can occur by chance, we generated 10,000 random sets, and counted the number of sets that resulted in a number of co-expressions equal or higher than the one observed. The random sets were generated by randomly scrambling the cells that express *myc* within each cluster of the 10xGenomics single-cell mRNA data.

2.6 Isolation of *S. domuncula* genomic DNA from adult wild-type sponge tissue

The sponge explants were lysed in 1.4 ml RLT plus lysis buffer (Qiagen RNeasy plus mini-Kit) containing 1% β- mercaptoethanol in 15 ml Falcon tubes. The explants in the lysis buffer were vortexed until the tissue was completely lysed. Afterwards, the lysates were stored at -80°C for at least 15 min or until processed.

For gDNA isolation 700 µl of lysate were passed through a gDNA eliminator column and the column was then washed with 500 µl Buffer AW1 and centrifuged for 30 sec at 10.000 rpm. Moreover, the DNA was washed with 500 µl Buffer AW2 and centrifuged for 1 min. After a final centrifugation step of the empty column for 1 min, the DNA was eluted in 100 µl EB buffer. Afterwards, the DNA samples were stored at -20°C.

2.7 Generation of *S. domuncula* cDNA from adult wild-type sponge tissue

The sponge explants were lysed in 1.4 ml RLT plus lysis buffer (Qiagen RNeasy plus mini-Kit) containing 1% β- mercaptoethanol in 15 ml Falcon tubes. The explants in the lysis buffer were vortexed until the tissue was completely lysed. Afterwards, the lysates were stored at -80°C for at least 15 min or until processed.

For RNA isolation 700 µl of lysate were passed through a gDNA eliminator column and 700 µL 70 % ethanol was added to the flow-through. Then the samples were transferred to an RNeasy spin column and centrifuged for 30 sec at 10.000 rpm. Afterwards, 350 µl Buffer RW1 was added at the top of the column, and centrifuged for 30 sec. To eliminate the remaining DNA in the RNA sample, 10 µl DNase I in 70 µl RDD Buffer were added directly to the column and incubated for 15 min at room temperature. Afterwards, 350 µl Buffer RW1 were added and the column was centrifuged again for 30 sec. Next, two washing steps with 500 µl Buffer RPE followed. Subsequently, the RNA was eluted in 38 µl RNase free water.

cDNA was then made from the RNA using the LunaScript™ RT SuperMix Kit (NEB). 4 µl of the RT-mix were added to 16 µl of RNA, incubetd for 2 min on room temperature and heated to 55°C for 10 min (thermoblock). After that the RT enzymes were inactivated by heating the samples to 95°C for 1 min. The cDNA was then stored at -20 °C.

2.8 Characterization of the *S. domuncula myc* locus

The sequence of the transcript identified as the unique S. domuncula *myc* (TR23970_c3_g1_i5) was used to design primers at the edge of the *myc* transcript, spanning the complete ORF. With those primers, the DNA segment containing the *myc* ORF was amplified via PCR from genomic DNA, and the amplicon was subsequently sequenced. The same primers were used to amplify the *myc* gene from a cDNA sample and the amplicon was subsequently sequenced. To identify possible introns within the *myc* locus, I compared the sequences resulting from the genomic and cDNA amplification. To identify potential SNPs within the *myc* locus, sequences from three S. domuncula specimens were analyzed.

2.9 Morpholino treatment

A vivo-morpholino (Gene tools, LLC) against the exon-intron boundary of the *S. domuncula myc* gene was designed by Gene Tools to inhibit splicing of the newly discovered *myc* intron (**Fig. 10**).

S. domuncula explants approximately 0.5 cm x 0.5 cm x 0.2cm in size were taken from an adult sponge and transferred into 24- well plates filled with 1 or 2 ml filtered seawater as described[6]. To treat explants with vivo-morpholino oligos, the standard procedure was to cut the explants, incubate them for 1h at room temperature in the 24-well plate and add vivo-morpholino to a final concentration of 1-20µM. After another incubation of 2h at room temperature the samples were put at 15°C for the respective incubation time.

To treat explant outgrowths with the vivo-Morpholino oligo, explants were cut and cultivated for 3 days in larger (ø6 cm) petri dishes with 8 ml filtered seawater. A small (8mm diameter) coverglass was placed under each explant, so that the explants would attach on the

coverglass. The explants forming outgrowth were transferred to a 24-well plate and treated with vivo-morpholino oligo as described above.

2.10 Simultaneous extraction of genomic DNA and total RNA from explants

The sponge explants were lysed in 1.4 ml RLT plus lysis buffer (Qiagen RNeasy plus mini-Kit) containing 1% β- mercaptoethanol in 15 ml Falcon tubes. The explants in the lysis buffer were vortexed until the tissue was completely lysed. Afterwards, the lysates were stored at -80°C until processed.

The genomic DNA (gDNA) and total RNA of the sponge explants were extracted using Qiagen's RNeasy kit. The lysates were vortexed for 30 sec and then centrifuged 5 min at 400 rpm at 4 °C. 700µl lysate were processed at once of which the total RNA was separated from the gDNA by centrifuging the lysate for 1 min at 10.000 rpm in the gDNA eliminator spin columns. First, the flow-through was used for the RNA extraction, while the gDNA in the gDNA eliminator spin columns was stored at room temperature.

700 µL 70 % ethanol was added to the flow-through. Then the samples were transferred to an RNeasy spin column and centrifuged for 30 sec at 10.000 rpm. Afterwards, 350 µl Buffer RW1 was added at the top of the column, and centrifuged for 30 sec. To eliminate the remaining DNA in the RNA sample, 10 µl DNase I in 70 µl RDD Buffer were added directly to the column and incubated for 15 min at room temperature. Afterwards, 350 µl Buffer RW1 were added and the column was centrifuged again for 30 sec. Next, two washing steps with 500 µl Buffer RPE followed. Subsequently, the RNA was eluted in 30 µl RNase free water.

During some of my experiments I needed to detect foreign *gfp* RNA in my cells and had therefore to completely eliminate the *gfp* DNA which was very abundant in the cells. During these experiments, therefore I performed additional DNase I digest steps. For this 10 µl RDD Buffer, 2.5 µl DNase I and 57.5 µl RNase free water were added to the RNA sample. Then the digest was performed for 15 min at room temperature. Afterwards, the sample plus 350 µl RLT plus lysis Buffer and 250 µl 100% Ethanol was loaded on an RNeasy spin column and centrifuged for 30 sec. Then one washing step with 750 µl Buffer RWI and two washing steps with 500 µl Buffer RPE followed. Again, the RNA was eluted in 30 µl RNase free water. This second DNase I digest was then repeated a second time including the consecutive washes.

Finally, the RNA sample was eluted in 36 µl RNase free water and either directly used for the synthesis of cDNA or stored at -80°C.

Afterwards, the sponge gDNA in the gDNA eliminator spin columns was isolated. For this 500 µl Buffer AW1 was added at the top of the DNA column and centrifuged for 30 sec at 10.000

rpm. Moreover, the DNA was washed with 500 µl Buffer AW2 and centrifuged for 1 min. After a final centrifugation step of the empty column for 1 min, the DNA was eluted in 100 µl EB buffer. Afterwards, the DNA samples were stored at -20°C.

2.11 Generation of cDNA from explants

Reverse transcription was performed on total RNA isolated from explants using the LunaScript™ RT SuperMix Kit (NEB). 4 µl of the RT-mix were added to 16 µl of RNA, incubetd for 2 min on room temperature and heated to 55°C for 10 min (thermoblock). After that the RT enzymes were inactivated by heating the samples to 95°C for 1 min.

2.12 qPCR analysis

Primers for qPCR were designed using the ’primer 3’ web page (http://bioinfo.ut.ee/primer3-0.4.0/) (Table 6). qPCR was performed on a qPCR machine (Quantstudio) using Luna® Universal qPCR Master Mix (NEB) as reporter dye. In all experiments, 10 µl Luna Master Mix and 1 µl primer mix (10 mM each primer) were pipetted into each well of the 96-well plate (Hard-Shell Low-Profile Thin-Wall 96-Well skirted PCR plate, Bio-Rad, Vienna, Austria). Afterwards, 0.4 µl of DNA or cDNA were added and filled up with RNase free water to a final volume of 20 µl per well. Furthermore, three technical replicates were used per sample. After an initial denaturation of 1 min at 95˚C, amplification and reading of the reporter dye were carried out at 60˚C for 30 sec over a total of 40 cycles with a 15 sec denaturation step at 95˚C at the beginning of each cycle. Finally, the melting curves of the amplified products were recorded. As reference gene *rpl11* was chosen because of its uniform expression observed previously[6].

2.13 Assembly of overexpression and dominant negative constructs

In order to express the desired proteins in the studied explants, two different constructs were assembled: Myc-overexpression construct and DN-Myc construct. For this, first the different parts were amplified, including the 3´UTR plus upstream flanking region, the desired *myc* sequence, the GFP sequence and the 5´ UTR plus flanking region.

These pieces were then combined using fusion PCR with primers overlapping 24nt into the next piece to be fused. Additionally, the overlapping primers between *myc* and GFP featured an additional sequence between the primer binding site and the overlap encoding the T2A site. These pieces were assembled in two rounds of fusion PCR, fusing 5´UTR with *myc* and *gfp* with 3´UTR during the first round and the two resulting pieces in the second round.

Fusion PCR reaction: 5 µl 10x LA PCR Buffer (Takara Bio, Otsu, Japan), 2.5 µl 10 mM dNTPs, 5 µl 25 mM MgCl2 (Takara Bio) , 2.5 µl primer left (10 µM), 2.5 µl primer right (10 µM) , 2.5 ng of each template, 0.25 µl LA Taq DNA Polymerase (Takara Bio), fill up with dH2O to 50 µl total reaction.

Fusion PCR program: 1) 95 °C, 5 min, 2) 94 °C, 45 sec, 3) 60 °C, 1 min, 4) 72 °C, 1 min/ kbp, 5) repeat step 2. to 4. 30 times, 6) 72 °C, 10 min, 7) 10 °C on hold. If the PCR product of a reaction was used inn another reaction 0.5 µl were taken from the unpurified PCR and then used as template.

The resulting construct were cloned into the pJET 1.2 vector. Since cloning into this vector necessitates that the insert has blunt ends, the inserts to be cloned were amplified by PCR using the Phusion High-Fidelity DNA Polymerase (NEB), which generates amplicons with blunt ends.

Blunting PCR reaction: 32.5 µl dH2O, 10 µl 5x HF Buffer (NEB), 1 µl dNTPs, 2.5 µl Primer L, 2.5 µl Primer R, 0.5 µl Phusion® High-Fidelity DNA Polymerase (NEB), 1 µl template PCR 1:10 diluted

Sequences of primers used for generation of the constructs:

Primer Database Number	Primer Name	Sequence
4871	94-F	GAGAACAACCCCTTGCTTCA
4873	160-R	CTACGCGATCGCTTTGTATAGT
4872	162-R	CACCAGCAATTTTGAAAGCA
4881	220-R	TGGATAAGAGATGAAAGCTCCTAAA
4898	243-F	ATGGGCTTCGAGCCTCTG
4899	244-R	CTCCACGTCCCCGCATGTTAGAAGACTTCCCCTGCCCTCTTAGCAAACGACTGGCATTG
4900	245-F	TGAACTATACAAAGCGATCGCGTAGATCCATTGTCCATCGCAAGT
4901	246-F	CCCAGAGCTGCTTCATTTTC
4902	247-R	CACAGGAAACAGCTATGAC
4903	248-R	GAGTCAGTGAGCGAGGAAGC
4904	249-F	AGGGAAGAAAGCGAAAGGAG
4905	250-F	TTTAGAGCTTGACGGGGAAA
4906	251-R	AGTTGGCAGAGGCTCGAAGCCCATGTCACACATTTTGGCTAGTTTTT
4874	252-F	CTTCTAACATGCGGGGACGTGGAGGAAAATCCCGGCCCAATGAGTAAAGGAGAAGAACTTT
4907	253-F	AATGAACACAAGGAAACAACG
4908	254-R	TGACGTTGTTTCCTTGTGTTCATTGTCACACATTTTGGCTAGTTTTT

2.14 DNA transfection

Chemical transfection with the polyethyleneimine-based transfection reagent JetPEI from the company Polyplus (France, Illkirch) was used to introduce the Myc-overexpression and DN-Myc constructs into the sponge tissue, similar to Revilla-i-Domingo et al., 2018 with minor modifications[6]. The transfection procedure was done in 24- well plates filled with 2 mL filtered seawater and 0.5 cm x 0.5 cm sponge explants were used. First, 4µg of DNA construct per sample were diluted in 100 µl of 150 mM NaCl (Polyplus) to a final volume of 100 µl per sample. In addition, the chosen amount polyethyleneimine (JetPEI) was diluted in 150 mM NaCl to a final volume of 100 µl per sample. Both solutions were mixed and briefly centrifuged. Then the polyethyleneimine solution was added to the DNA solution and immediately vortexed and centrifuged. In addition, I performed a "no transfection" control, where the polyethyleneimine was omitted, meaning that the DNA was diluted in a final volume of 200 µl of 150 mM NaCl. The polyethyleneimine + DNA solution was incubated for 30 min at room temperature. Finally, I added 200 µl of the polyethyleneimine + DNA solution per well to the sponge explants. I incubated the plate gently rocking for 3 to 4 hours at room temperature. The samples were then returned to 15 °C.

2.15 shRNA design

shRNAs were designed using invivoGens siWizard software (http://www.invivogen.com/sirnawizard/index.php). The motif size was set to 19 and from the resulting sequences DNA sense and antisense templates for shRNA synthesis were made according to the scheme in **Table** 1. For each gene three shRNAs were designed this way to increase silencing and all of them were analyzed using mfold from the University of Albany (http://unafold.rna.albany.edu/?q=mfold) to minimize alternative secondary structures.

2.16 shRNA production via in vitro transcription

The shRNAs were made following the protocol of Karabulut et al.[37]. First the 0.5µl of the desired F and R templates (100mM) were mixed in 57 µl of nuclease free water and heated to 95°C for 5 min to remove secondary structures. The templates were then cooled to room temperature. Meanwhile a mastermix of 8 µl10X RNAPol Buffer (NEB), 8 µl Ribonucleotides (final concentration of 2mM), 2µl Ribolock RNase inhibitor (NEB) and 4 µl SP6 polymerase (NEB, 20 U/µl) per sample was prepared and then added to the templates. The in vitro transcription reaction was incubated for 4h at 40°C. After this the RNA was purified using the Direct-zol™ RNA MiniPrep kit according to the manufacturer's RNA purification protocol with DNase I digest. The RNA concentration was then measured via Nanodrop to an average result of ~ 20ng/µl.

2.17 shRNA transfection

For shRNA transfection non-liposomal cationic amphiphile transfection reagent INTERFERin from the company polyplus (France, Illkirch) was used. To do this, pieces of sponge tissue (0.5 cm x 0.5 cm x 0.2cm in size) were dissociated into single cells by pressing explant sized pieces of tissue through a 35 µm nylon mesh cell strainer (1 explant per ml). Then those cells were transferred to a 24-well plate, 500µl per well in filter sea water and incubated at room temperature for 1h before treatment. Meanwhile the INTERFERin shRNA complexation was prepared. 3 different shRNAs per gene were diluted to a final concentration of 20nM each in the same solution with either 150nM NaCl or DMEM medium to a final volume of 100 µl per well. Then the INTERFERin was vortexed and 2µl were added for each 100 µl. The solution was then vortexed immediately for multiple seconds and incubated at room temperature for 10 minutes. Then 100 µl of the solution were added to the 500 µl dissociation. The samples were then incubated at room temperature for 3h before being returned to 15°C and were incubated at 15°C for another 45h.

Table 1

Template sequences for shRNA synthesis. In the schematic row the exemplary template consists of the **SP6 promoter sequence**, **the sense motif**, **the loop region**, **the antisense motif** as well as a **3' TT overhang**.

Primer Database Number	Schematic	**ATTTAGGTGACACTATAG** **GAGAGATCCAGTCTGACAT** **TTCAAGAGA** **ATGTCAGACTGGATCTCTC** **TT**
4929	Myc 42-F	ATTTAGGTGACACTATAG GCACATCAACTCAAGTCCA TTCAAGAGA TGGACTTGAGTTGATGTGC TT
4930	Myc 42-R	AA GCACATCAACTCAAGTCCA TCTCTTGAA TGGACTTGAGTTGATGTGC CTATAGTGTCACCTAAAT
4931	Myc 1661-F	ATTTAGGTGACACTATAG GACGAGCTTCGCACAACAT TTCAAGAGA ATGTTGTGCGAAGCTCGTC TT
4932	Myc 1661-R	AA GACGAGCTTCGCACAACAT TCTCTTGAA ATGTTGTGCGAAGCTCGTC CTATAGTGTCACCTAAAT
4933	Myc 1770-F	ATTTAGGTGACACTATAG GAAGGTGACAATCTTGAGA TTCAAGAGA TCTCAAGATTGTCACCTTC TT
4934	Myc 1770-R	AA GAAGGTGACAATCTTGAGA TCTCTTGAA TCTCAAGATTGTCACCTTC CTATAGTGTCACCTAAAT
4935	SILICATEIN 190-F	ATTTAGGTGACACTATAG GAGAGATCCAGTCTGACAT TTCAAGAGA ATGTCAGACTGGATCTCTC TT
4936	SILICATEIN 190-R	AA GAGAGATCCAGTCTGACAT TCTCTTGAA ATGTCAGACTGGATCTCTC CTATAGTGTCACCTAAAT

4937	SILICATEIN 225-F	ATTTAGGTGACACTATAG GACGATAGTCAGGGAGAGA TTCAAGAGA TCTCTCCCTGACTATCGTC TT
4938	SILICATEIN 225-R	AA GACGATAGTCAGGGAGAGA TCTCTTGAA TCTCTCCCTGACTATCGTC CTATAGTGTCACCTAAAT
4939	SILICATEIN 257-F	ATTTAGGTGACACTATAG GGCAGAGGAGGAAGATGAA TTCAAGAGA TTCATCTTCCTCCTCTGCC TT
4940	SILICATEIN 257-R	AA GGCAGAGGAGGAAGATGAA TCTCTTGAA TTCATCTTCCTCCTCTGCC CTATAGTGTCACCTAAAT

3. Results

3.1. Identification of the *S. domuncula myc* and *max* genes

To obtain the sequence of the *S. domuncula myc* transcript I performed a *tblastn* search of the mouse c-Myc protein sequence against a previously established *Suberites domuncula* transcriptome. Every transcript with an E-value smaller then 1e-4 was taken and further analyzed by blasting against other genomes (*Mus musculus, Ephydatia fluviatilis* and *Amphimedon queenslandica*). All transcripts that did not return a Myc protein as a first hit in the blast analysis, or transcripts with other obvious problems, like sequence fusions, were discarded. In this way, I found a single bona fide candidate *S. domuncula myc* transcript: transcript TR23970_c3_g1_i5. Since most animals possess multiple *myc* genes (including the cnidarian *Hydra vulgaris*[28]*),* I looked at the presence of *myc* genes in a different sponge species, for which the transcriptome is available (*Ephydatia fluviatilis*) and in the genome of one sponge species for which the genome sequence is available (*Amphimedon queenslandica*). Consistent with the presence of a single *myc* gene in the sponge phylum, in both cases only one *myc* gene was found, as in *S. domuncula.* From this analysis, I conclude that sponges contain a single *myc* gene, and the multiple *myc* genes in other animals is the result of gene duplications after the evolutionary split between sponges and other animals.

To confirm that the identified transcript (TR23970_c3_g1_i5) is indeed a *myc* gene, I proceeded to construct a phylogenetic tree comparing Myc proteins from different species. As an outgroup, I used the protein Max, as this is the most similar protein to Myc in the mouse proteome. Myc and Max bHLH domains dimerize and show strong similarities. The *S. domuncula max* sequence I obtained in the same way as *myc* and also found a single copy with the transcript ID TR21690|c0_g3_i4. The presence of a *max* gene in *S. domuncula* is in line with findings in the other sponge species (see below).

I found that the Myc protein sequence obtained from transcript TR23970_c3_g1_i5 clusters with the Myc sequences of other animals (**Fig. 5**). This analysis confirms that transcript TR23970_c3_g1_i5 indeed encodes for *S. domuncula* Myc. The obtained phylogenetic tree is also consistent with a single *myc* gene in sponges, and gene duplications that occurred after the split between sponges and other animals.

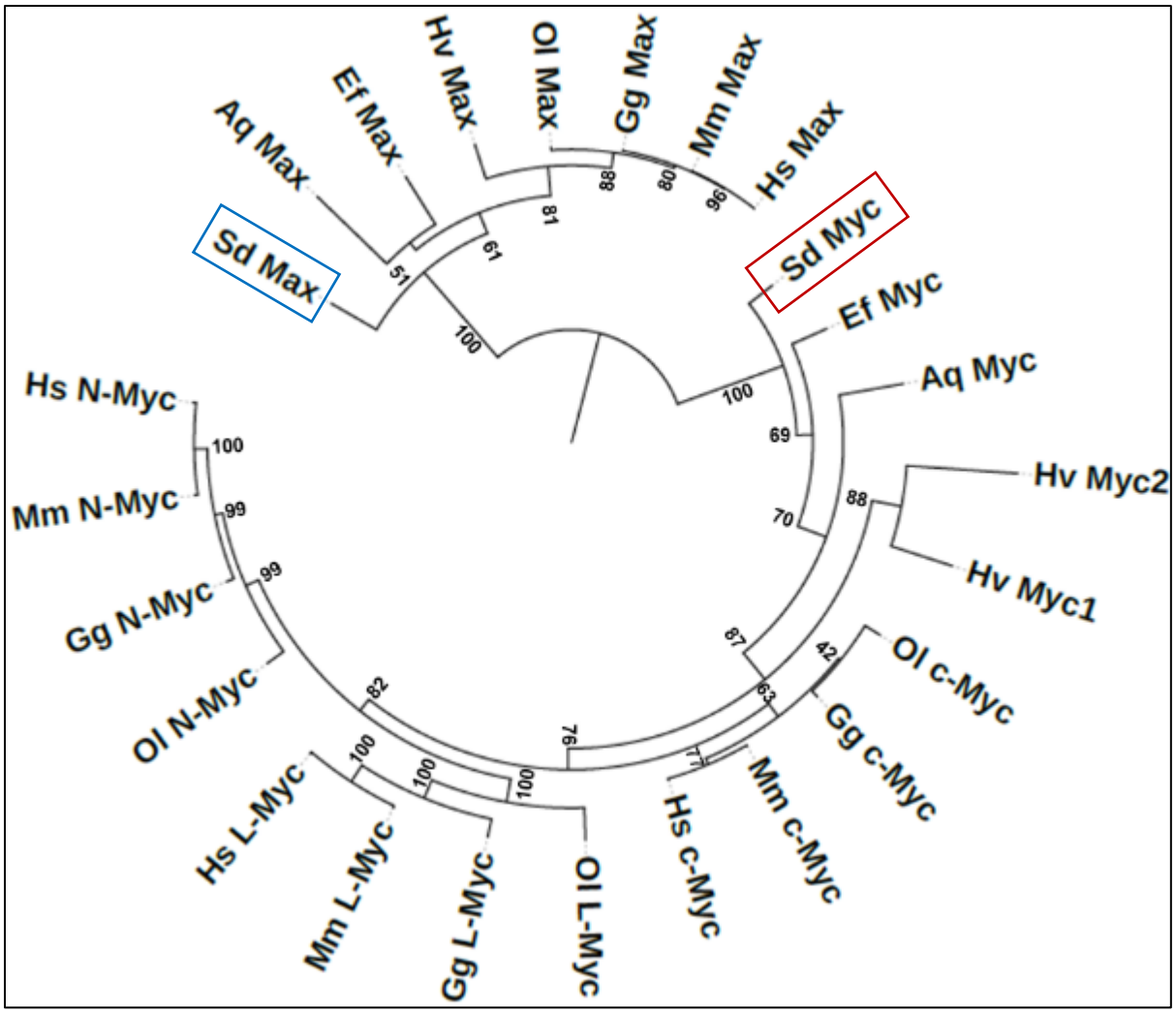


Figure 5: Phylogenetic tree of Myc and Max proteins.

The phylogenetic tree confirms the identity of *S. domuncula* Myc, and it is consistent with a single Myc protein in sponges. Sd= *Suberites domuncula*, Ef= *Ephydatia fluvatilis*, Aq= *Amphimedon queenslandica*, Hv= *Hydra vulgaris*, Ol= *Oryzias latipes,* Gg= *Gallus gallus*, Mm= *Mus musculus*, Hs= *Homo sapiens*

Further analysis using protein domain prediction programs like Pfam or InterPro showed that the *S. domuncula* Myc protein has two conserved protein domains. A Myc N-terminal domain and a basic helix-loop-helix domain, which are also found in all Myc proteins of all examined organisms (**Fig. 6**). This structure coincides with the protein structure observed for Myc in other sponges analyzed (*A. queenslandica* and *E. fluviatilis*), as well as with that of mouse N-Myc (**Fig. 6**). In addition to these two domains, mouse c-Myc shows an additional domain, a leucine zipper domain, which is completely absent from the other mouse Myc proteins and from all the sponge Myc proteins. Mouse L-Myc proteins show a duplication of the N-terminal domain while the N-Myc proteins have the same protein domains as the sponge Myc proteins. Analyzing Max proteins, the same way showed that the *S. domuncula* as well as the mouse Max proteins both only have a single bHLH domain.

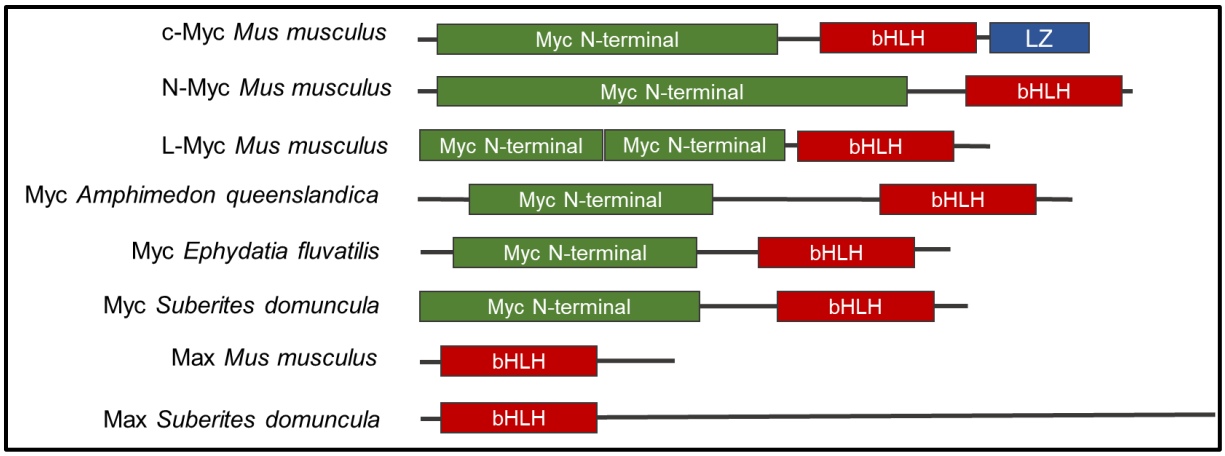


Figure 6: Structural domains of different Myc proteins.

Sponge Myc proteins have the same functional domains as vertebrate Myc proteins, with the exception of an additional leucine zipper domain in vertebrate c-Myc. The *S. domuncula* Max protein also has the same functional domain as the mouse Max protein.

This, in addition to the clustering in the phylogenetic tree **(Fig. 5)** strongly suggests that *S. domuncula* has only one copy of *myc* as well as *max*, and the transcripts that corresponds to these genes are TR23970_c3_g1_i5 for *myc* and TR21690|c0_g3_i4 for *max*.

3.2. Investigation of conservation of direct Myc targets

As mentioned in the introduction, a recent study reported 100 genes found to be directly regulated by Myc in a human cell line [32]. Since this is, to date, the most exhaustive, unbiased list of direct targets of Myc, we took advantage of this list to test, in an unbiased way, whether these direct targets might unravel an ancestral Myc-regulated network. To this end, I analyzed the top 50 of these 100 genes by looking for the homologous genes in *Suberites domuncula* (**Table 2**), and found 36 homologs of those 50 genes. If any of these genes is a direct transcriptional target of Myc in *S. domuncula*, we expect that this gene should be co-expressed with *myc* at the level of individual cells. Exploiting the availability of single-cell mRNA-seq data in *S. domuncula* (**Fig. 7A**), and knowing the identity of the *S. domuncula myc* gene, I tested whether any of the 36 candidate genes is statistically significantly co-expressed with *myc* (see Materials and Methods). I found 17 genes significantly co-expressed with *myc* ($p>0.05$) (**Table 2**). The overlap between the expression of Myc and that of these 17 genes can be visually appreciated (see **Fig. 7B,C** for one example). To be sure that the calculated statistical significance values do not represent artifacts of our method, we calculated co-expression statistical significance values for a number of genes that are not associated to Myc in any way (negative controls), and found that none of those resulted in significant p-values (see **Table 3** for a few examples).

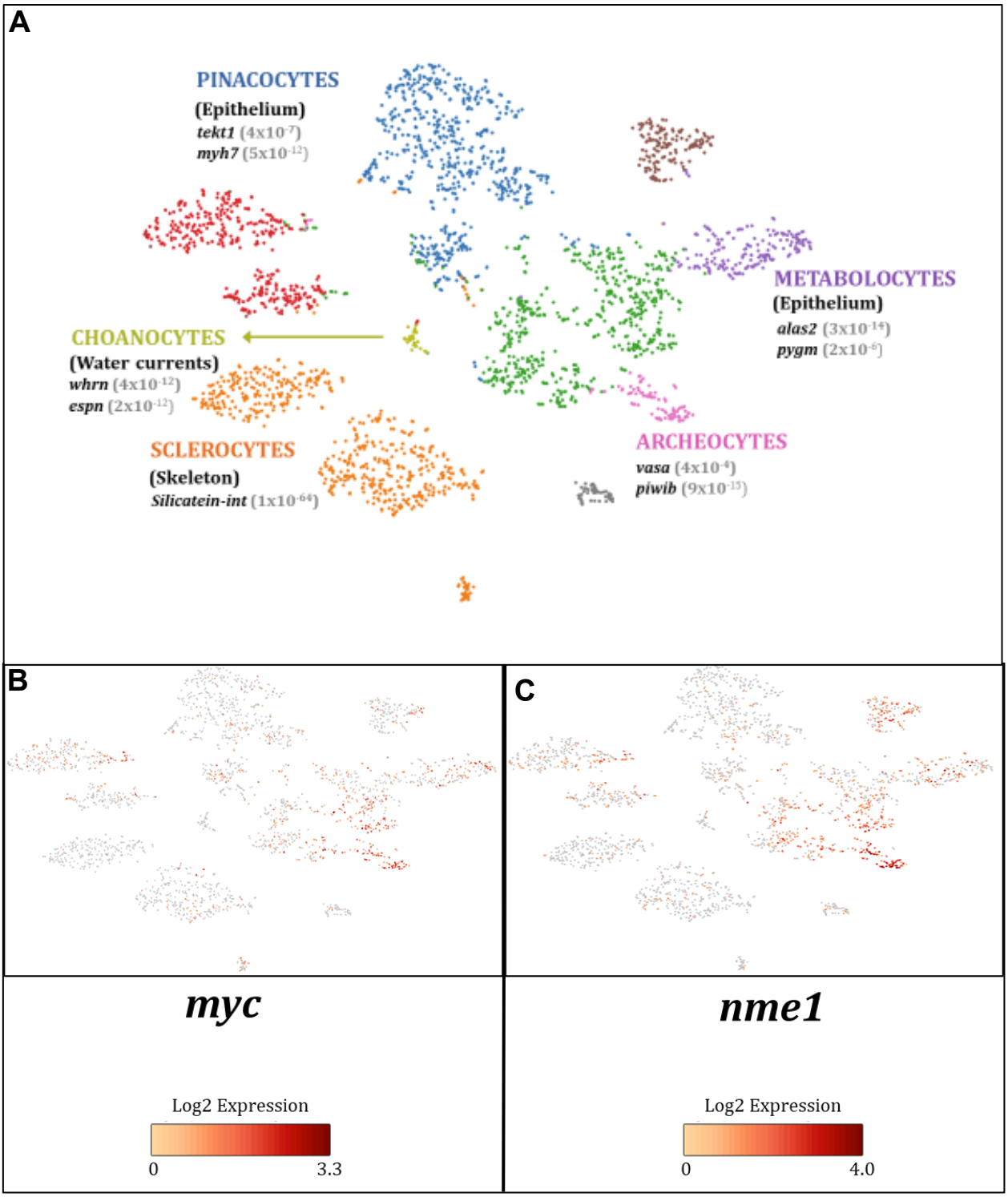


Figure 7: single cell data visualization

A) t-SNE plot of *S. domuncula* single-cell mRNA-seq data.
B) Expression levels of *myc* in the cells of the t-SNE plot of **Fig. 7A**
C) Expression levels of *nme1* in the cells of the t-SNE plot of **Fig. 7A**

Table 2

50 of the top 100 genes found to directly be regulated by Myc[32].

EntrezID (Mouse)	Mouse gene symbol	E value *S. domuncula* homolog search	*S. domuncula* homolog ID	p-value Myc co-expression	Function
79080	CCDC86	3,00E-08	TR53916|c0_g1_i1	1,00E-04	NA
10849	CD3EAP	-	-		
79709	COLGALT1	5,00E-134	TR23941|c1_g1_i5	0,1481	
9188	DDX21	0.0	TR10938|c0_g1_i3	0,0043	Ribosome biogenesis
56919	DHX33	2,00E-142	TR37708|c2_g1_i3	0,0236	Ribosome biogenesis
27292	DIMT1	1,00E-136	TR18696|c0_g1_i3	0,0241	Ribosome biogenesis
1802	DPH2	3,00E-19	TR38037|c1_g1_i1	1	
10969	EBNA1BP2	3,00E-52	TR19706|c0_g1_i5	0,3408	
84650	EBPL	5,00E-54	TR28434|c0_g1_i1	1	
2171	FABP5	-	-		
2194	FASN	1,00E-21	TR16786|c0_g1_i1	1	
24147	FJX1	8,00E-31	TR52344|c0_g1_i1	0,8693	
51083	GAL	-	-		
26354	GNL3	6,00E-47	TR37202|c0_g1_i1	0,3242	
3099	HK2	0.0	TR37743|c2_g1_i1	1,00E-04	Glycolysis
3614	IMPDH1	0.0	TR44466|c1_g1_i5	1,00E-04	Nucleotide synthesis
3939	LDHA	-	-		
3945	LDHB	-	-		
55646	LYAR	1,00E-34	TR10198|c0_g1_i2	0,0097	Ribosome biogenesis
10514	MYBBP1A	-	-		
9603	NFE2L3	1,00E-14	TR48134|c1_g1_i1	1,00E-04	NA
84365	NIFK	3,00E-26	TR17100|c0_g1_i1	1,00E-04	Ribosome biogenesis
4830	NME1	4,00E-76	TR23057|c0_g1_i1	1,00E-04	Nucleotide synthesis
654364	NME1-NME2	-	-		
4831	NME2	2,00E-74	TR23057|c0_g1_i1	1,00E-04	Nucleotide synthesis
9221	NOLC1	5,00E-32	TR48063|c2_g1_i4	1,00E-04	Ribosome biogenesis
51491	NOP16	9,00E-08	TR23971|c0_g1_i2	0,0053	Ribosome biogenesis
10528	NOP56	0.0	TR53010|c0_g2_i3	0,5107	
441478	NRARP	-	-		
10606	PAICS	-	-		
201164	PLD6	4,00E-07	TR7299|c0_g1_i1	1	
84172	POLR1B	0.0	TR35971|c0_g1_i3	0,0203	Ribosome biogenesis
5464	PPA1	1,00E-104	TR24027|c0_g1_i1	0,1936	
5471	PPAT	-	-		
6004	RGS16	4,00E-17	TR52410|c2_g1_i8	0,2504	
9045	RPL14	3,00E-32	TR37050|c0_g1_i2	1,00E-04	Ribosome biogenesis
23212	RRS1	2,00E-58	TR34433|c0_g1_i5	1,00E-04	Ribosome biogenesis
26156	RSL1D1	2,00E-35	TR15490|c1_g1_i1	1,00E-04	Apoptosis

150094	SIK1	1,00E-105	TR21020|c0_g1_i1	0,3476	
102724428	SIK1B	3,00E-120	TR24923|c1_g2_i23	1	
58516	SINHCAF	-	-		
6566	SLC16A1	-	-		
285958	SNHG15	-	-		
6723	SRM	1,00E-118	TR36569|c2_g1_i6	0,0929	
7023	TFAP4	2,00E-23	TR19725|c0_g1_i3	0,0086	Cell cycle
56652	TWNK	4,00E-26	TR24191|c0_g1_i1	0,8278	
7371	UCK2	1,00E-70	TR37668|c0_g2_i3	1	
10885	WDR3	0.0	TR33912|c1_g1_i4	1	
8187	ZNF239	5,00E-41	TR37668|c0_g2_i3	1	
162979	ZNF296	7,00E-09	TR17603|c0_g3_i3	0,2364	

I next looked at the known function of the human homologs of the 17 *S. domuncula* genes found to be significantly co-expressed with *myc* (**Table 2**). 9 of these 17 genes are associated to ribosome biogenesis, 3 in nucleotide synthesis, 1 in glycolysis, 1 in cell cycle progression and 1 in apoptosis. For 2 of these 17 genes, their function is not known. Interestingly, these functions are directly linked to known Myc-regulated stem cell functions: ribosome biogenesis and nucleotide synthesis are directly linked to cell proliferation, an essential function of stem cells, while glycolysis has been shown to control stem cell differentiation in mouse hematopoietic stem cells[38–40], mouse spermatogonial stem cells[41], and Drosophila neuromasts[41]. While the direct link between Myc and stem cell proliferation has been extensively documented[32,41,42],some studies have also shown that control of glycolysis by Myc is responsible for Mycs control of stem cell fate[41].

In summary, these results provide a list of interesting candidate direct targets of Myc in *S. domuncula*, which, if confirmed experimentally, would allow us to point at a conserved Myc-regulated network with relevance to stem cell functions. Given that no techniques exist to address gene function in sponges, I next intended to establish a method to modulate *myc* expression in *S. domuncula*, with the intention to test the role of *myc* in sponge stem cells, as well as to experimentally test the regulation of our candidate targets of Myc.

Table 3: Negative control genes for statistical significance test of *myc* co-expression.

Gene name / transcript ID	p-value
Actb (expressed everywhere)	0.4
TR24037|c0_g1_i3 (Archeocyte cluster)	0.6
Whrn (Choanocyte cluster)	0.8
Silicatein (Sclerocyte cluster)	0.9

3.3. Testing vivo-Morpholino oligos to knock down *myc* function

Next, I wanted to manipulate *myc* expression levels to monitor the effect this would have on the sponge stem cells. To do this I established a technique to treat sponge explants with antisense vivo-morpholino oligos. Morpholinos are oligomer molecules with DNA bases attached to methylenemorpholine rings that are connected via phosphorodiamidate groups. They have a similar structure to RNA and are able to basepair with complementary sequences but will not be degraded by RNAses or DNAses because of their different backbone, making them very stable *in vivo*. The “vivo” in vivo-morpholino refers to an octoguanidine dendrimer modification of these morpholinos that facilitates uptake by treated cells[43].

In other mode systems, vivo-Morpholino oligos can be used to inhibit translation or splicing[44]. Inhibition of splicing has the advantage that it allows assessing the functionality of the vivo-Morpholino oligo by assessing the proportion of spliced vs non-spliced transcripts by PCR[44–46]. Therefore, I decided to check whether the *S. domuncula myc* gene contains an intron that would allow knocking down the function of the gene by targeting a vivo-Morpholino against splicing.

3.3.1. Characterization of the *S. domuncula myc* locus

To check whether the *S. domuncula myc* gene contains an intron, I used the sequence of identified *myc* transcript (TR23970_c3_g1_i5) to design primers that spanned the entire *myc* ORF. I then used these primers to amplify DNA from (i) genomic DNA, and (ii) cDNA. Comparison of the PCR amplicons obtained from the genomic DNA and cDNA samples revealed that the *S. domuncula myc* gene contains an intron (**Fig. 8**). The genomic DNA PCR amplicon was then sequenced, and revealed a single 332bp intron.

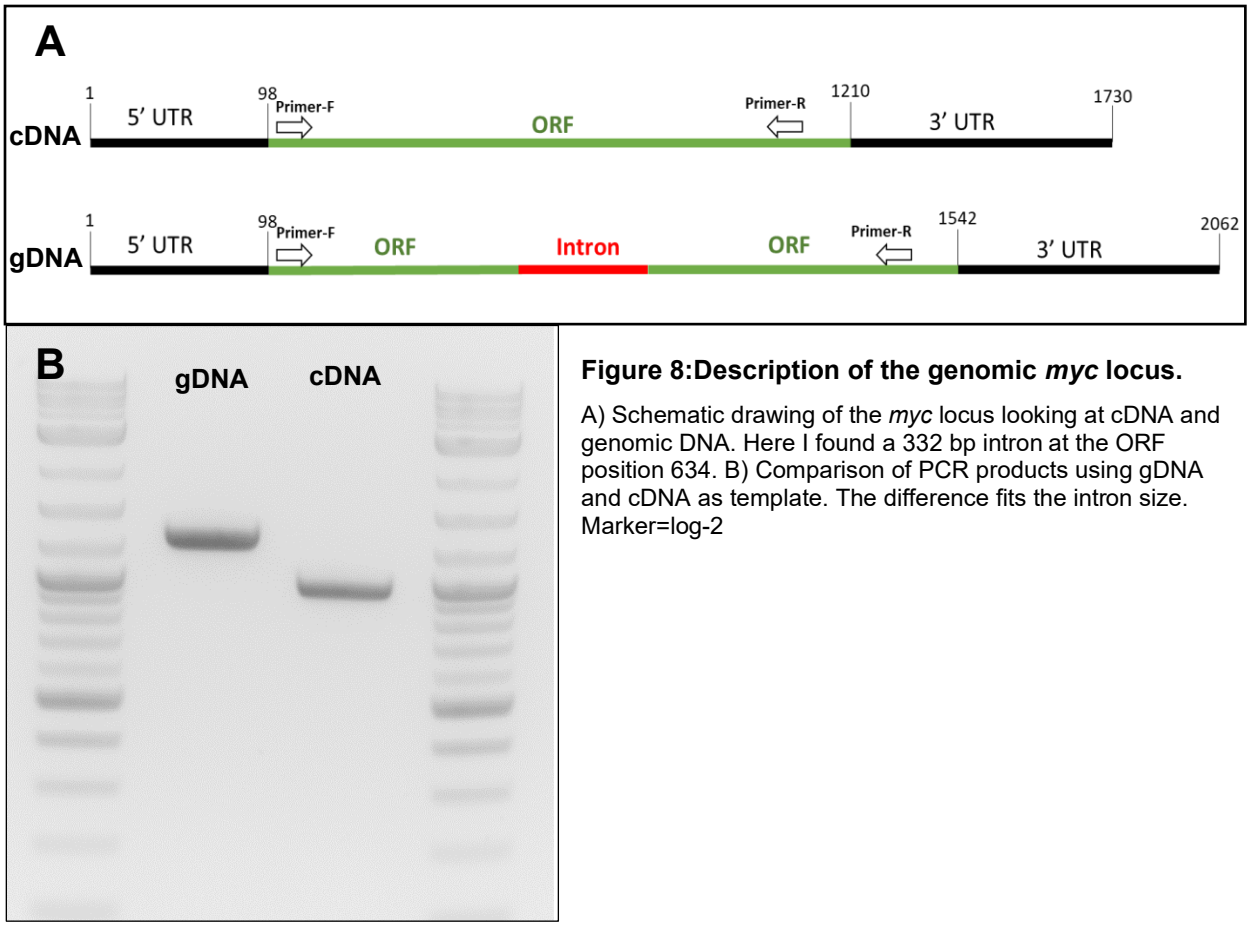


Figure 8:Description of the genomic *myc* locus.

A) Schematic drawing of the *myc* locus looking at cDNA and genomic DNA. Here I found a 332 bp intron at the ORF position 634. B) Comparison of PCR products using gDNA and cDNA as template. The difference fits the intron size. Marker=log-2

To further check whether inhibition of splicing of the *myc* intron would be a suitable strategy to knockdown Myc function, I inspected the sequence of the intron, and found that inhibition of splicing would result in a premature STOP codon (**Fig. 9A**). Further analysis of the protein sequence revealed that the inhibition of *myc* splicing would result in a Myc protein that would lack the bHLH domain and part of the transactivating domain (**Fig. 9B**), leaving it unable to bind to Max, rendering it therefore non-fun ctional[29].

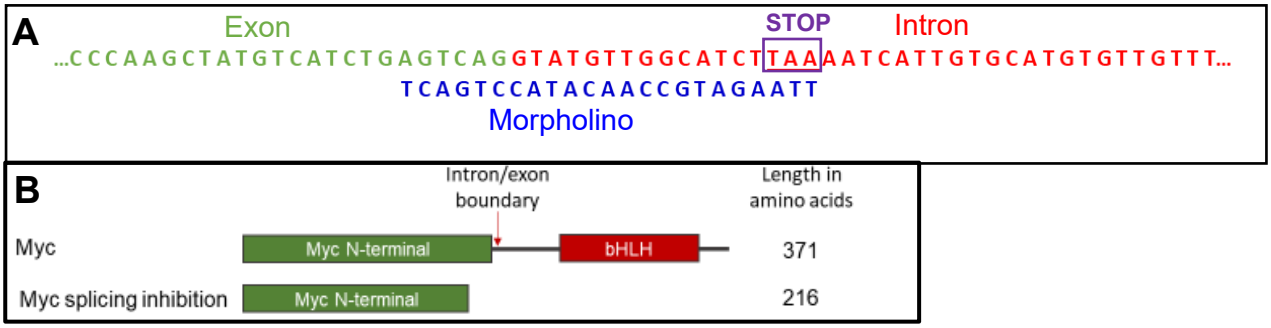


Figure 9A: Schematic of the designed morpholino binding at the intron/exon junction. **B:** Schematic of the functional protein compared to the truncated protein after splicing inhibition.

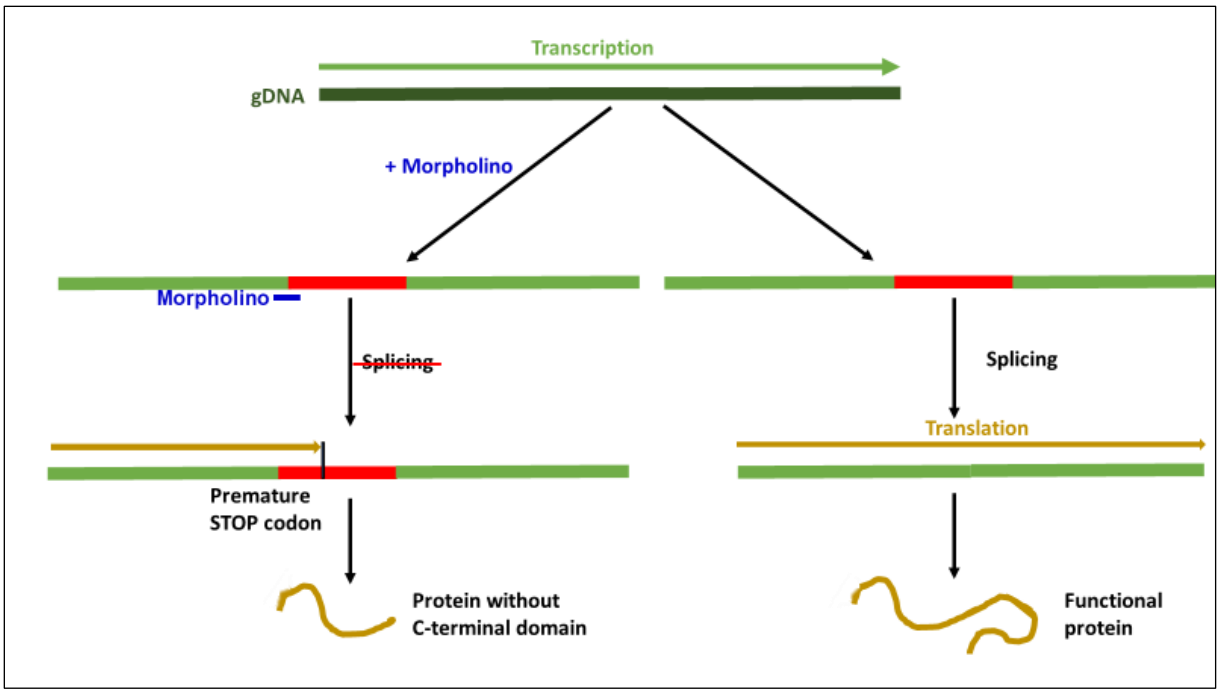


Figure 10: Schematic of morpholino working mechanism.

The binding of the morpholino to the 3' splice site of the mRNA should mask the sequence for the splicing machinery, therefore failing to process the transcript correctly. Incomplete processing would then lead to a premature STOP codon in the intron, cutting the protein short and rendering it non-functional.

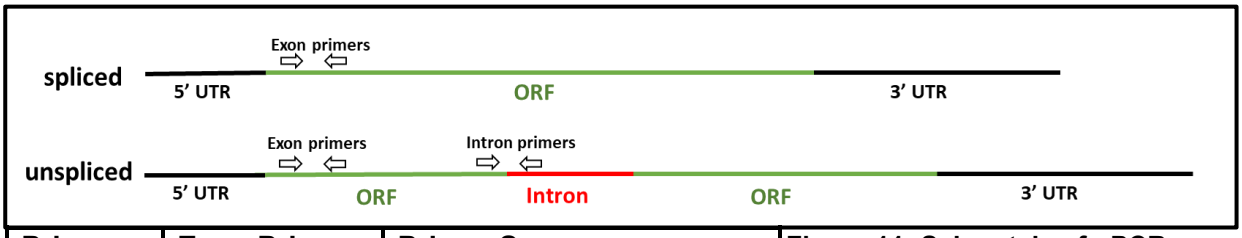
Based on this characterization, I decided to design a vivo-Morpholino that targets *myc* splicing (**Fig. 10**). To be sure that the region targeted by the intended vivo-Morpholino (the exon-intron boundary), does not have polymorphisms that could jeopardize the function of the vivo-Morpholino oligo, I sequenced the region around the intron/exon boundary in three different *S. domuncula* specimens. All three contained an intron as described above, and no polymorphisms were detected in the region around the exon-intron boundary. Therefore, I submitted the target sequence to the GeneTools company, and the company designed a vivo-Morpholino oligo that spans the Myc exon-intron boundary (**Fig. 9A**).

3.3.2. Establishment of a sensitive assay to detect inhibition of splicing

Given the size difference of the *myc* transcript with or without the intron (**Fig. 8B**), PCR amplification of the *myc* sequence from a cDNA sample followed by gel electrophoresis is a suitable approach to detect the inhibition of splicing. However, given its non-quantitative character, this approach is not suitable to detect very small levels of splicing inhibition. Therefore, I devised a quantitative assay that would be able to detect even small levels of splicing inhibition, based on quantitative PCR (qPCR).

To establish such a quantitative assay, I designed qPCR primers within the *myc* intron (*myc* intron primers) and within one of the exons (*myc* exon primers) **(Fig. 11)**. I then used these qPCR primers to measure the levels of *myc* exon and *myc* intron in a sample of *S.*

domuncula cDNA, obtained by reverse transcription of a total RNA sample. To control for the presence of genomic DNA in the detection of *myc* exon and intron, I also ran a qPCR on a "NoRT Control" sample, consisting of a total RNA sample where the reverse transcription reaction had been run without the reverse transcriptase enzyme. This quantification revealed the presence of *myc* intron sequences within an untreated RNA sample (**Fig. 12**, green bars), reflecting small levels of nascent transcripts present in these samples.



Primer database numbers	Exon Primers	Primer Sequence
4896	237-F	TGACAGTCCAACGACCAATC
4897	238-R	CAGACTCCGTGGTTCGTACA
	Intron primers	
4891	231-F	CCAGATCATCCACCTCTGCT
4893	234-R	CTCCCACAAACAACACATGC

Figure 11: Schemtaic of qPCR primer binding sites and primer sequences.

To be able to confirm that this assay would be able to quantitatively detect even small levels of splicing inhibition, I measured the level of *myc* exon and *myc* intron in samples where known amounts of intron-containing *myc* sequences (obtained by amplification of the *myc* gene from a genomic DNA sample) were added to the cDNA sample (**Fig. 12**). This experiment showed that even when the amount of intron-containing *myc* sequences was as low as 1% of the total amount of *myc* sequences in the cDNA samples, our assay was able to quantitatively detect these extra intron sequences.

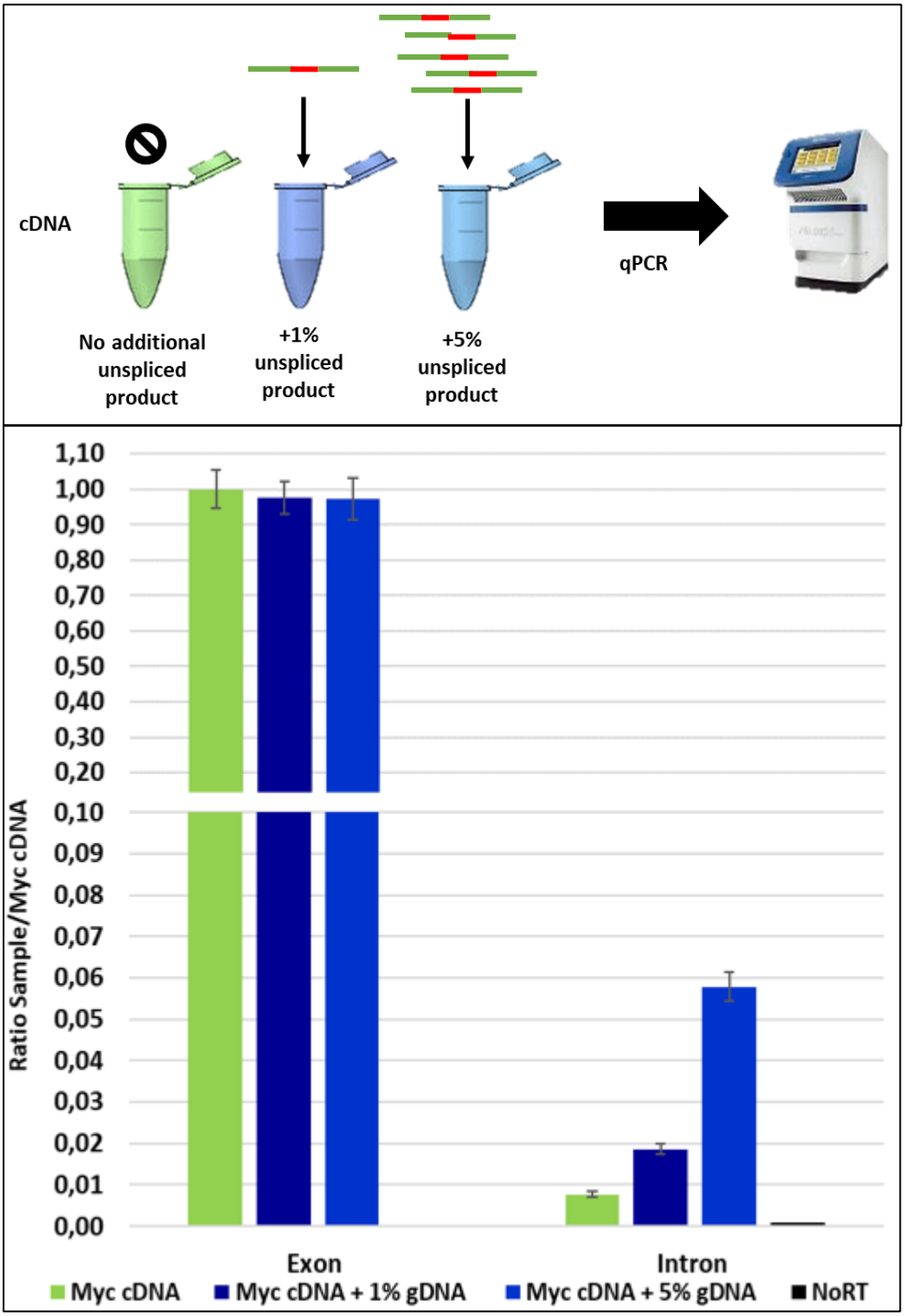


Figure 12: Testing the sensitivity of our assay to detect slicing inhibition.

Compared next to each other here is cDNA from the same sample with different amounts of *myc* intron sequences added. What can be observed here is that there is a basic level of unspliced *myc* transcript approximately equal to 0.6% of total *myc* transcripts,. This test demonstrates that our assay is capable of detecting even 1% of splicing inhibition.

3.3.3. Morpholino treatment

Having a sensitive assay to detect splicing inhibition, I proceeded to test whether treatment with our vivo-morpholino oligo against *myc* splicing would result in any increase of *myc* intron sequences.

To deliver the vivo-Morpholino oligos, I cut explants from an adult sponge as described in Materials and Methods, and placed them in a 24 well plate. As explained above, vivo-morpholinos (Gene Tools LLC.USA), have a covalently linked delivery moiety. Therefore, to deliver it into the cells of the sponge explants, I added the vivo-morpholino oligo (up to final concentrations ranging from 1μM to 20μM) into the seawater containing the explants, and incubated the explants for various lengths of time, ranging from 6 to 48 hours (**Fig. 13A**). After the desired incubation time, total RNA was extracted from treated and untreated samples. cDNA was generated by reverse transcription and qPCR was performed using *myc* exon and *myc* intron primers (**Fig. 13A**).

The ratio of *myc* intron over *myc* exon sequences in treated samples was calculated and compared to untreated samples **(Fig. 13B-E**). Inhibition of splicing should result in a higher *myc* intron over *myc* exon ratio in the treated samples compared to untreated samples, which was not observed in any of the conditions tested (**Fig. 13B-E**). Treatment of explants with 10μM vivo-Morpholino oligo for 24 hours resulted in a significant change in the *myc* intron over *myc* exon levels (**Fig. 13B**). However, since this represented an increase in the level of splicing, rather than inhibition of splicing, and the effect was not observed in any other condition, I did into further purse this result.

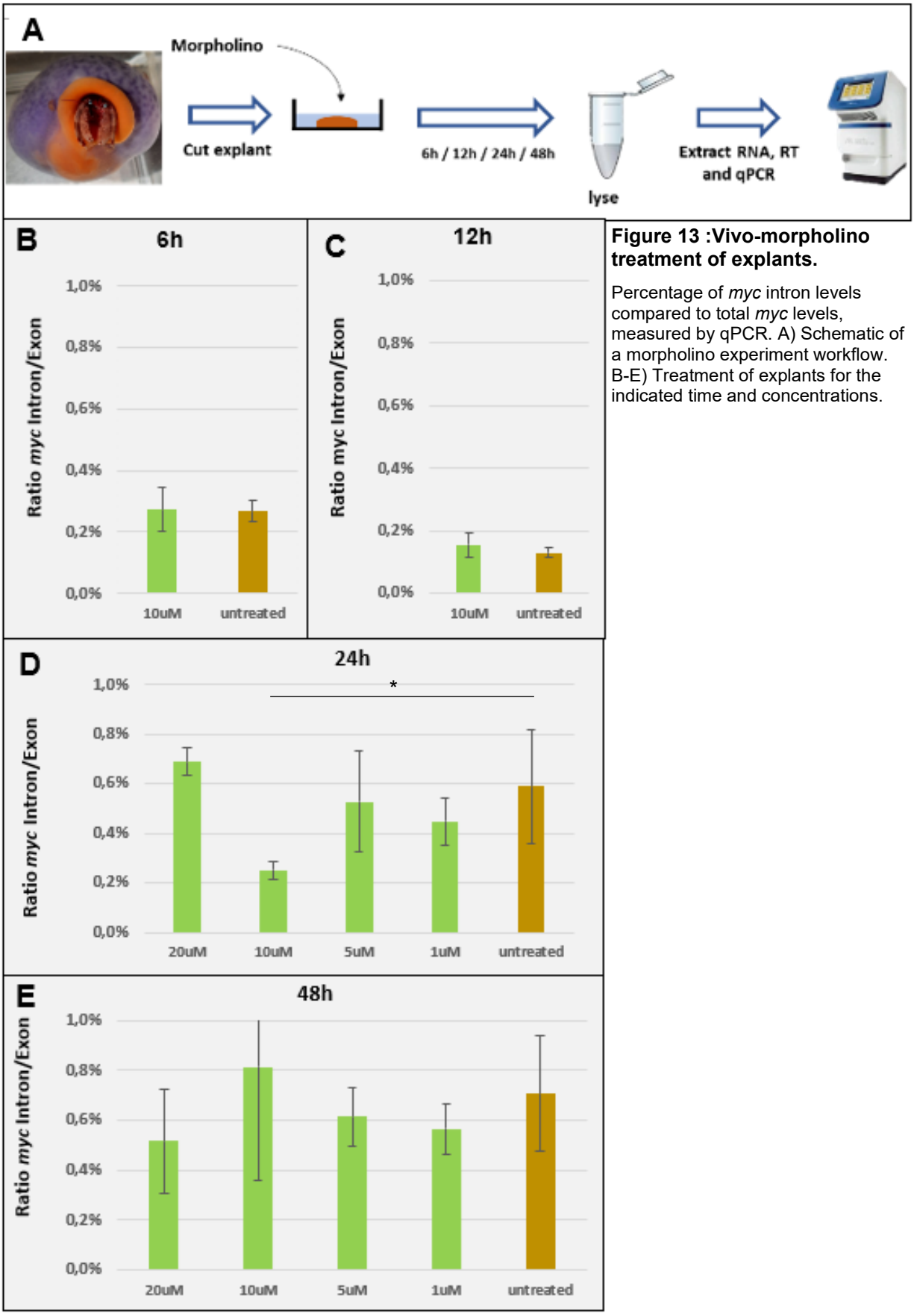


Figure 13 :Vivo-morpholino treatment of explants.

Percentage of *myc* intron levels compared to total *myc* levels, measured by qPCR. A) Schematic of a morpholino experiment workflow. B-E) Treatment of explants for the indicated time and concentrations.

One possible explanation for the lack of splicing inhibition by the vivo-morpholino is that the vivo-Morpholino is not able to penetrate into the sponge tissue. I therefore decided to test the vivo-Morpholino treatment in conditions that might facilitate the penetration of the vivo-Morpholino oligos.

S. domuncula explants are able to regenerate in 24-well plates[6]. In the first stage of regeneration, a thin (blastema-like) film of tissue is generated. Alluding to the fact that this thin tissue might be easier to penetrate than the thicker adult tissue, I tested the vivo-morpholino treatment on early regenerating sponge tissue (**Fig. 14A**; see Materials and Methods). Unfortunately, no change in unspliced *myc* mRNA levels could be detected in this condition either (**Fig. 14B**).

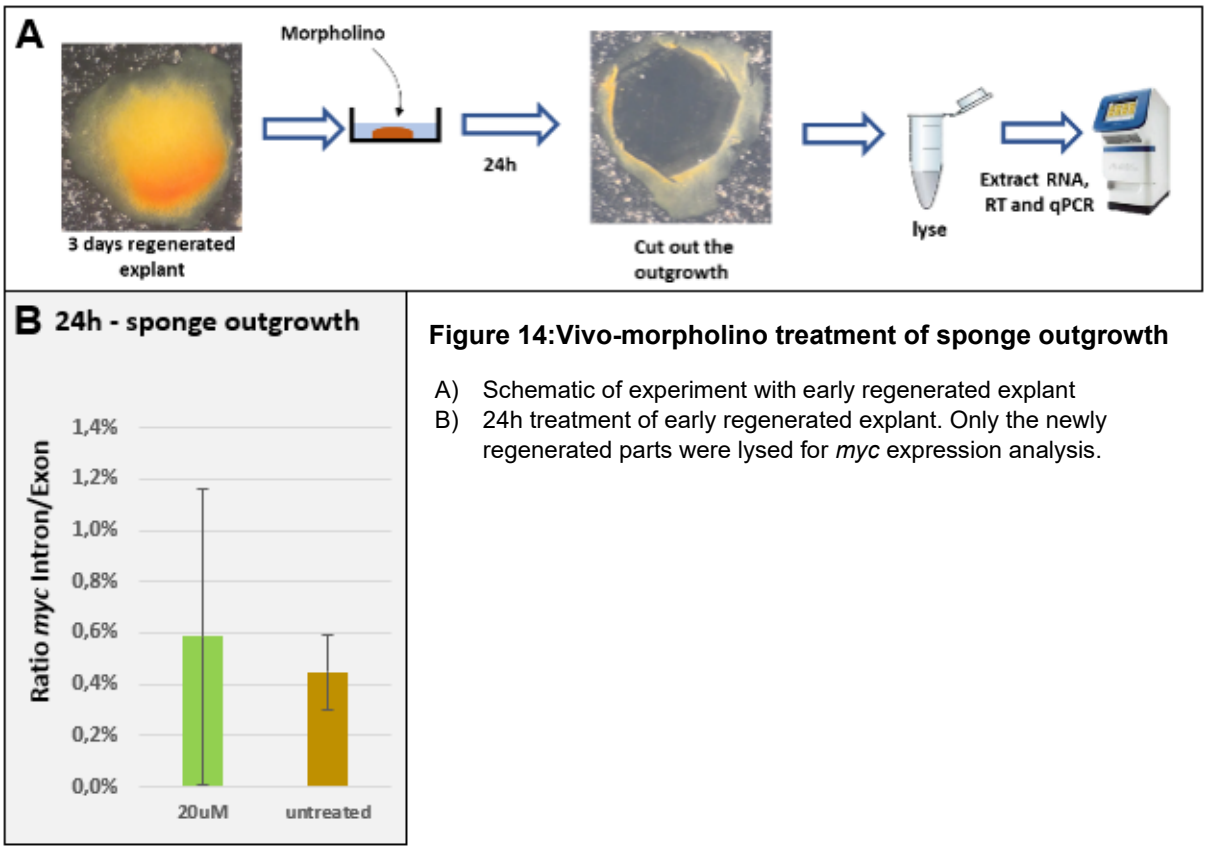


Figure 14:Vivo-morpholino treatment of sponge outgrowth

A) Schematic of experiment with early regenerated explant
B) 24h treatment of early regenerated explant. Only the newly regenerated parts were lysed for *myc* expression analysis.

In summary, I did not find any condition that resulted in inhibition of *myc* splicing by using the designed vivo-Morpholino. Therefore, I decided to test alternative techniques to modulate the function of *myc*.

3.4. Myc overexpression and expression of a dominant-negative form of Myc

Alternative ways to test Myc function include transfecting foreign DNA into sponge cells that result either in overexpression of Myc or expression of a dominant-negative Myc (dn-Myc) variant.

Previously, a construct was established that expresses Gfp under the control of the *S. domucnula actb* regulatory region[6]. *actb* is a highly abundant gene expressed everywhere in the sponge body[6]. This construct will from now on be referred to as Actb::Gfp. It has already been shown that Actb::Gfp can be transfected efficiently into sponge cells and that Gfp expression is detectable. Building on this approach, I intended to express either Myc or dn-Myc under the control of the same *actb* regulatory region.

3.4.1. Construction of overexpression and dominant-negative *myc* constructs

Starting from the Actb::Gfp construct, I made two constructs, one designed to express Myc and the other designed to express dn-Myc. The idea for the dn-Myc construct is that the dn-Myc form competes with wild-type Myc for DNA target sequences, while being incapable of activating transcription. To make such a dn-Myc construct, I removed, from the Myc coding sequence, the entire Myc N-terminal trans activating domain (TAD). The TAD is where the effector domains, also called Myc boxes, are situated. Missing these regions leaves Myc unable to enact its role as transcription factor[42], while still able to dimerize with Max and bind its target sequences[47]. Both construct were designed to express Gfp (in addition to the *myc* or *dn-myc* genes), so that the successful expression from the constructs could be monitored by qPCR against the *gfp* gene, or by fluorescence..

For the construction of both the Myc-overexpression construct, the 5' and 3' flanking regions as well as the *gfp* sequence of the Actb::GFP construct were amplified via PCR and assembled using fusion PCR (**Fig. 15A**). The *myc* and *gfp* sequences were set in frame so that bi-cistronic mRNA could be made. This had the advantage of later being able to detect Gfp via microscopy, identifying the cells expressing the foreign DNA. Additionally, a T2A ribosomal skip site was added between the *myc* and *gfp* genes, so that two separate proteins, Myc and Gfp, result from translation of the generated transcript [48]. Producing two separate proteins avoids potential undesired interference of GFP on the Myc protein, since there is no evidence that a C-terminally added Gfp sequence would not perturb Mycs function.

The dn-Myc construct was assembled in a similar way (**Fig. 15B**), except that the first 259 amino acids were deleted from the Myc sequence.

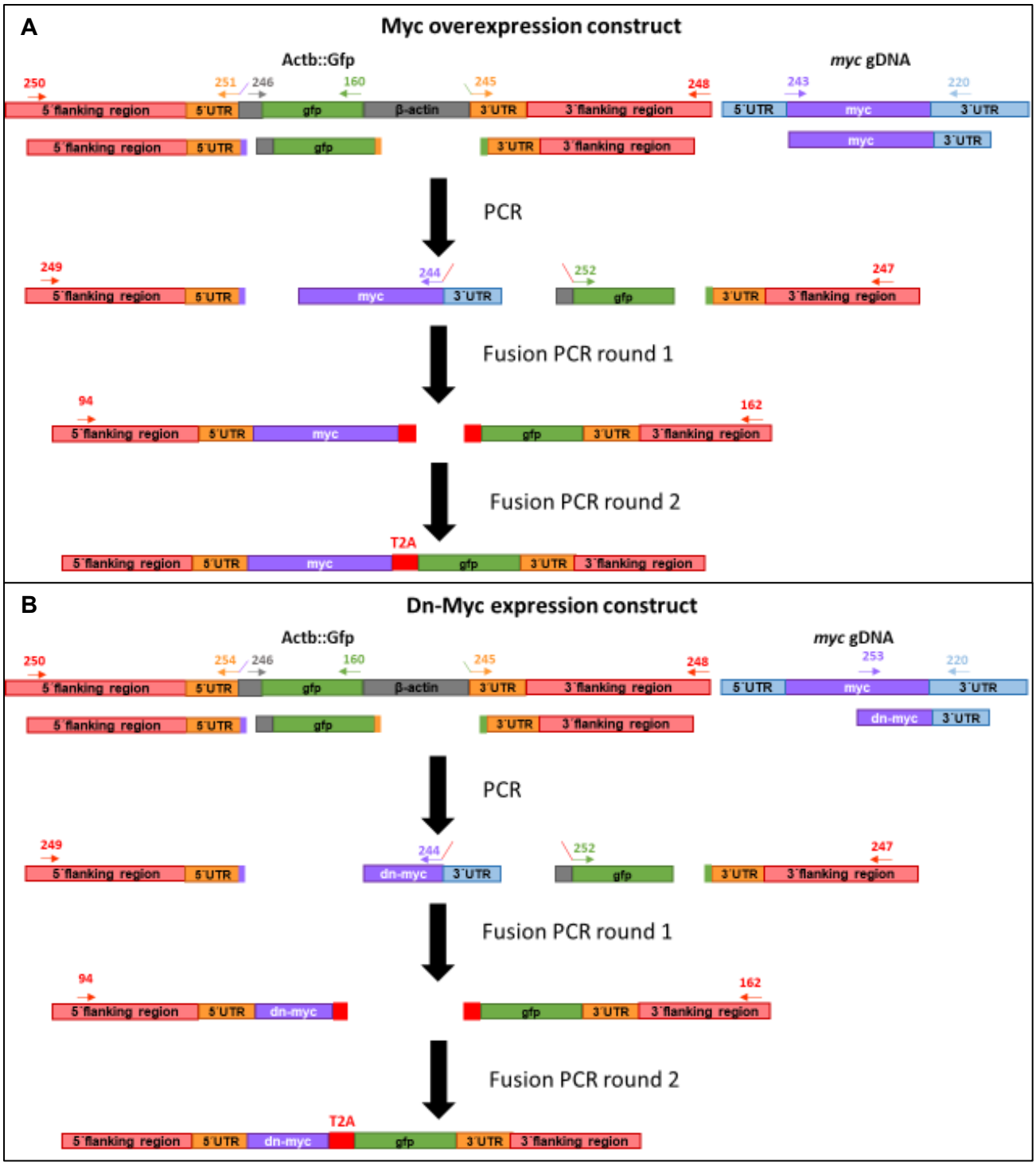


Figure 15: Illustration of the assembly of the overexpression (A) and dominant negative Myc construct (B).

All the depicted PCR steps were performed using LA-Taq DNA polymerase which adds a T-overhang to the product. Since the final product was ligated into a blunt vector (pJET 1.2) the product was blunted by another round of PCR using Phusion® High-Fidelity DNA Polymerase (not depicted).

3.4.2. Transfection of the Myc overexpression and dn-Myc constructs

The construct DNA delivery method I used was chemical transfection using a polyethyleneimine-based DNA transfection reagent (JetPEI, Polyplus) previously established for *Suberites domuncula* explants[6]. Standard size explants were prepared and incubated as described in Materials and Methods. To measure the number of transfected construct molecules per cell, genomic DNA was extracted from the transfected samples and the number of *gfp* molecules relative to a single copy gene in the genomic DNA samples was quantified by qPCR (**Fig. 16A**). As a positive control, I used the unmodified Actb::Gfp, for which transfection efficiency was already tested[49].

The results show a high transfection efficiency (**Fig. 16B**). Control samples, which were exposed to the exogenous DNA, but without the presence of the JetPEI transfected reagent ("-JetPEI"), showed no transfected DNA (**Fig. 16B**), confirming that the exogenous DNA detected in the +JetPEI samples was transfected inside the cells, and not, for example, floating in the medium.

Even though the biological variability was high, the transfections worked efficiently, ranging from ~25 copies per cell to ~450 copies per cell (**Fig 16B**). This even exceeds the previously reported transfection efficiency of 60-120 copies per cell[49].

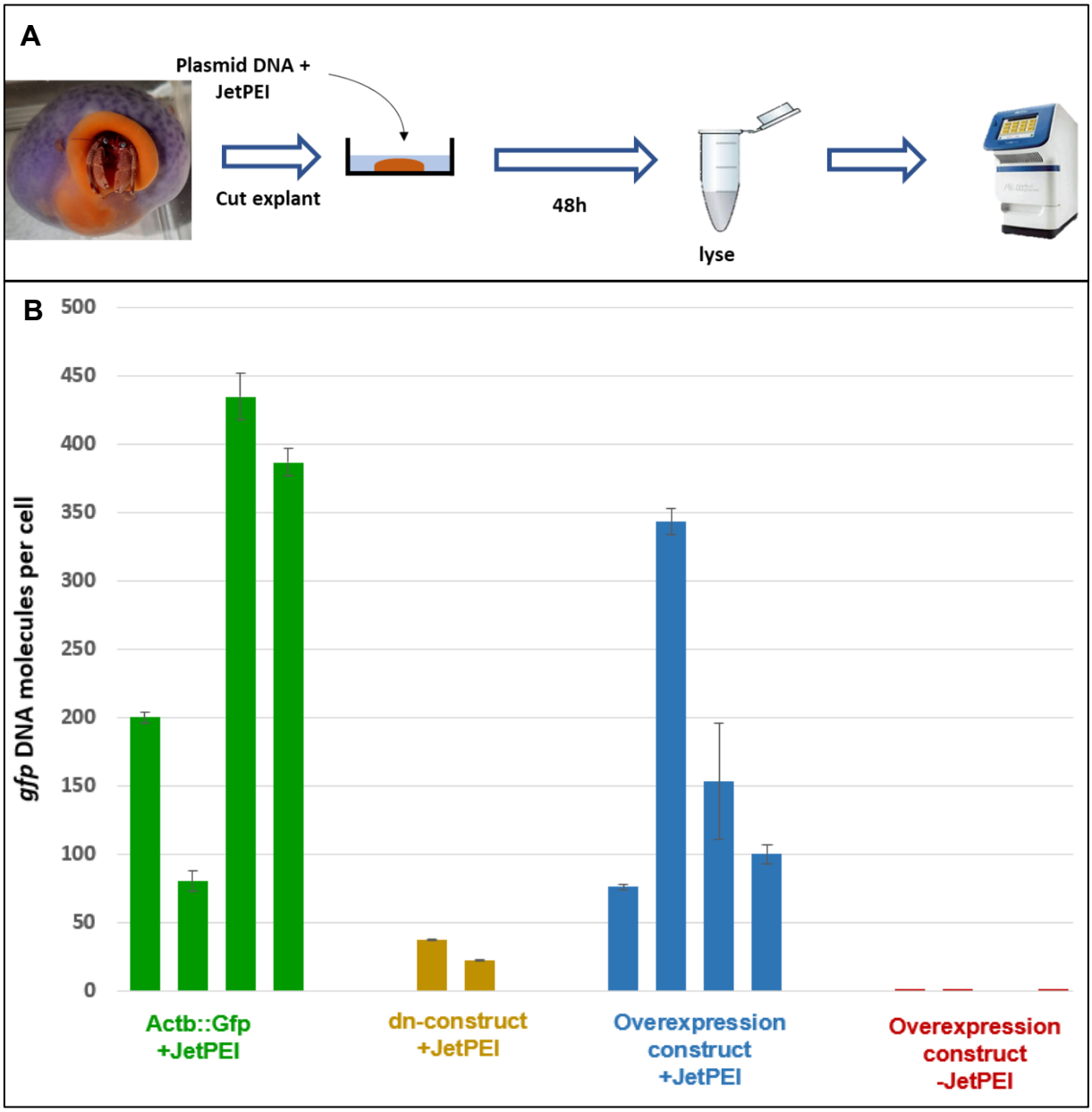


Figure 16: Plasmid transfection efficiency

A) Schematic of transfection experiment setup. In all instances 4µg of DNA were transfected using 12µl JetPEI. B) Transfection of constructs with JetPEI. In addition the overexpression construct was also transfected without JetPEI. Clearly without JetPEI no *gfp* DNA is detected, suggesting that the detected *gfp* DNA is indeed in the cells and not in the medium. Even though the plasmid copies per cell of the dn-Myc construct are vastly different than the others this is not unusual. The differences in transfection efficiency were subject to high biological variability with an unknown cause.

3.4.3. Detecting transcription from the Myc overexpression and dn-Myc expression constructs

Next, I analyzed the amount of transcription that resulted from the transfected constructs. For this, I took advantage of the presence of the *gfp* gene in both constructs. I extracted total RNA from the transfected samples, and generated cDNA. The level of *gfp* mRNA relative to a reference gene (*rpl11*[6]) was measured by qPCR by using the same *gfp* primers as for the DNA.

I was able to detect *gfp* transcripts in the Actb::Gfp samples (**Fig. 17**), at approximately the same levels as published[6,49]. In the Master thesis of C. Schmidt another reference gene (*g7a*) was used whose abundance is roughly 10% of *rpl11.* Therefore, the previously found *gfp* mRNA level of ~1% of *g7a* corresponds very nicely to the here found 0.15% of *rpl11* (please adapt the name of the genes to the genes mentioned in the paper). For the Myc-overexpression and dn-Myc constructs, no *gfp* mRNA could be detected (data not shown), which might be the result of multiple factors (see discussion).

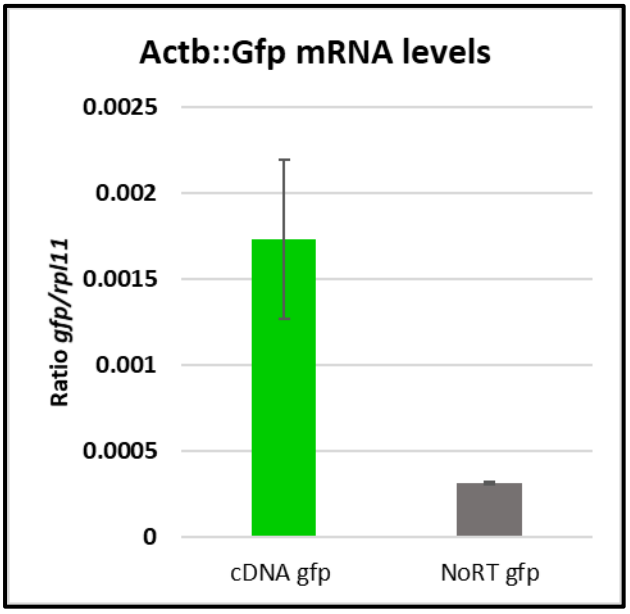


Figure 17: Transcription efficiency of the Actb::Gfp construct.

Measured *gfp* mRNA levels compared to *ribol11* mRNA levels. Even though very low the cDNA levels are clearly higher than their NoRT counterparts.

3.5. Testing shRNA to knock down Myc

As another way to approach gene knockdown, I decided to explore small hairpin RNA (shRNA) mediated RNA interference. RNA interference (RNAi) is another established way to knock down gene expression[50]. It works by introducing double stranded RNA (dsRNA) into cells, which can be achieved in two ways: transcribing it *in vivo* or transfecting it chemically. Inside the cells, the dsRNA then gets processed and loaded into the RNA-induced silencing complex (RISC). This complex then recognizes mRNA complementary to the processed RNA and cleaves it. Therefore, if the shRNA treatment is effective, we should see reduced

levels of the target gene in our treated samples. To deliver the shRNA, we decided to use INTERFERin, a transfection reagent designed for siRNA transfection in cell culture.

3.5.1. shRNA design

For desing and production of the shRNAs I followed the protocol of Karabulut et al.[37] (see Materials and Methods). I designed 3 shRNAs targeting the *myc* gene. In order to be able to distinguish between specific and non-specific effects, I also designed 3 shRNAs that targeted a gene that is expressed in a different cell type than *myc*: *silicatein.* s*ilicatein* is expressed in sclerocytes, and it encodes for a key enzyme in developing spicules, the main part of the siliceous skeleton of sponges.

3.5.2 shRNA production

shRNAs were synthesized from PAGE purified synthesized DNA oligos (**Table** 1) and from these aligned oligos the shRNAs were *in vitro* transcribed using T6 polymerase. The shRNA oligos were then purified using the Direct-zol™ RNA MiniPrep kit as recommended by the manufacturer and the amount of purified shRNAs was then determined via Nanodrop.

3.5.3. shRNA transfection and gene knockdown efficiency

The shRNAs were transfected using INTERFERin (Polyplus, France), which is a non-liposomal cationic amphiphile transfection reagent, designed for transfecting small RNAs. For this experiment, the sponge explants were dissociated into single cells (**Fig. 18**; see Materials and Methods) because INTERFERin was designed for cell culture conditions and I wanted to mimic these conditions as much as possible. The transfection was done for 48h before lysis, RNA extraction and reverse transcription. The mRNA levels of the target genes were analyzed using qPCR. No significant difference in mRNA levels could be detected between treated and untreated samples (**Fig. 18**).

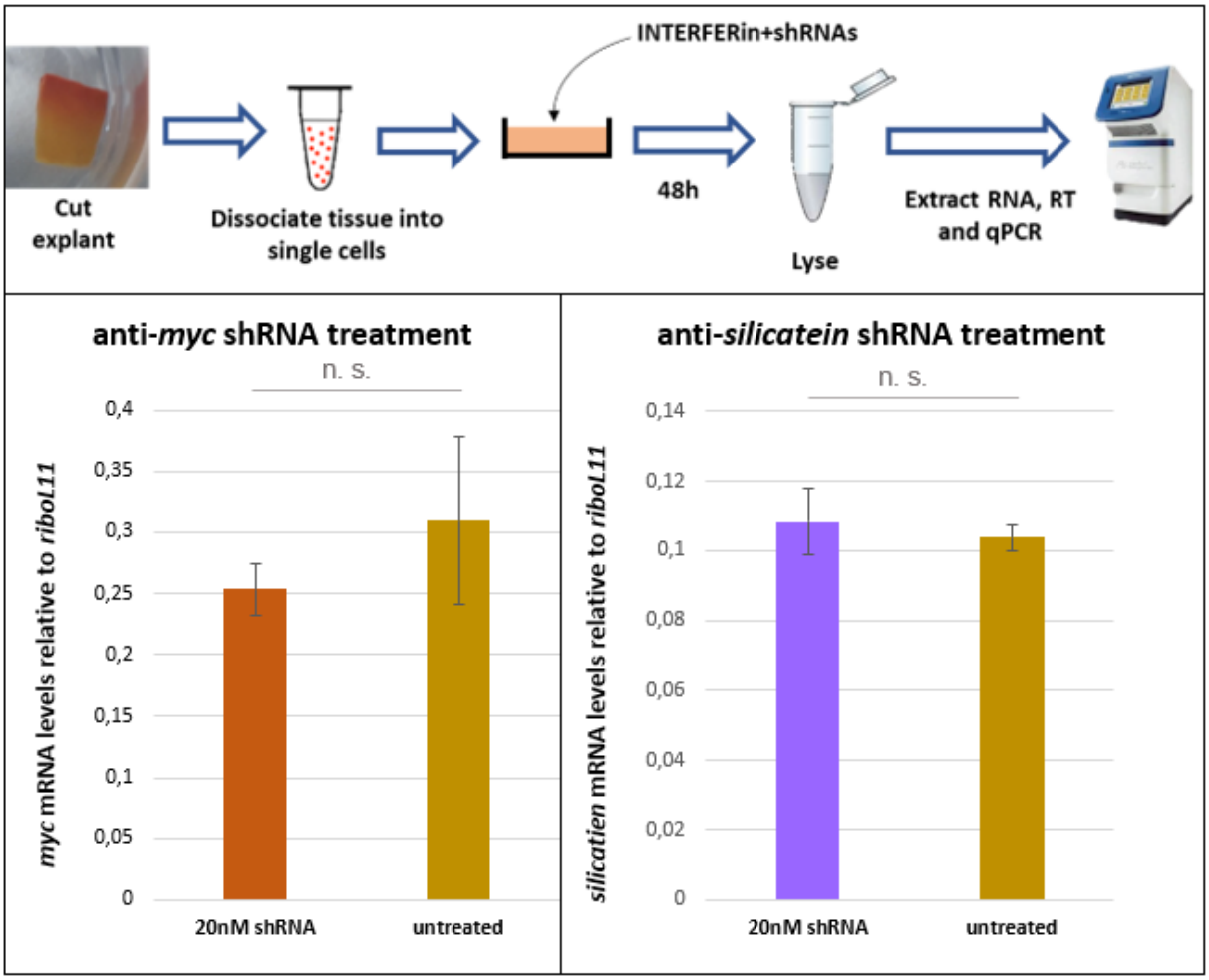


Figure 18 :mRNA levels after shRNA treatment

mRNA levels of *myc* and *silicatein* treated with respective shRNAs vs an untreated sample. No effect could be observed. Statistical analysis (t-test two sided) was performed for both treatments but no significant difference could be detected.

4.Discussion

As outlined in the Results, my overall work has contributed towards the general aim of assessing a potential conservation of Myc in stem cell regulation. Here, I am providing a digest of my results, grouped by the two main goals I followed.

1. To test conservation of direct targets of Myc in an unbiased way.

As presented in the Results, my analysis of direct targets of human Myc in leukemic cells showed that 36 out of 50 of those genes had a homologous gene in *Suberites domuncula*. Of these, 17 were co-expressed with *myc* with statistical significance, which is significantly higher than we would have expected if there was no conserved co-expression between *myc* and the investigated genes. We also tested the specificity of our analysis method by selecting genes that were expressed everywhere (*actb*) or in specific cell types (see Table 3). Most interesting here is TR24037|c0_g1_i3 since this transcript is specific to archaeocytes. In our analysis we found that *myc* and TR24037|c0_g1_i3 were not co-expressed, showing that significant co-expression of *myc* is not due to its association with the archaeocyte cluster.

Importantly, and as described in detail in the results section, the known function of the aforementioned mouse or human homologs of these 17 *S. domuncula* genes can be directly linked to specific cellular functions, including cell proliferation (e.g. *nme1, nme2, tfap4*), ribosome biogenesis (e.g. *rpl14, nifk, nolc1*) and glycolysis (*hk2*). This supports our hypothesis of a strongly conserved cellular function of Myc, since Myc has been shown to be a regulator of glycolysis in mamals[24,41] as well as a driver of proliferation[30,31,33]. Given the unique phylogenetic position of sponges, as one of the earliest branching metazoan phylum (see introduction), these results suggest a conserved Myc-controlled network for the regulation of stem cell functions in early animals.

These results are significant at two levels: A. They are consistent with existing hypotheses that suggest that Myc may have played a key role in the transition to multicellular life from unicellular ancestors [18,35,51] ; B. They provide a list of interesting candidate genes to look at to further explore the suggested ancestral Myc-regulated network.

2. Experimentally test the function of Myc in sponge stem cells.

During our identification of the *S. domuncula myc* and *max* we found that *S. domuncula* only contains single copies of both genes. This was an important finding since most of the data on Myc and Max stems from mammals which have multiple copies of Myc and other early branching animals like Hydra vulgaris also have more than one *myc* gene. For Max in mouse and *Hydra* only a single copy is present. Comparing our result to other sponges like

Ephydatia fluvatilis and *Amphimedon queenslandica,* for both of which either the genome or the transcriptome is known, we found that this appears to be common at least within the Demospongiae. The phylogenetic tree (**Fig. 5**) supports these findings since all the porifera Myc sequences cluster together, away from all other organisms. This suggests that in the early metazoans only one *myc* gene was present and that the other Myc families arose from this one variant after the last common ancestor of sponges and other animals.

As shown in the results, I approached the challenge of testing the function of Myc in sponge stem cells with 4 different methods, well established in other organisms: antisense vivo-morpholino oligos, Myc overexpression, dn-Myc expression and RNAi using shRNAs. While I did not achieve a significant down-regulation of Myc function, each of the experimental approaches generated insight that could be useful for improving the respective method:

My efforts to establish a method for *myc* splicing inhibition and subsequent gene knockdown using vivo-morpholinos were not successful. However, I was able to established an assay to accurately measure inhibition of *myc* splicing, which can be used in the future for follow-up experiments. One possible explanation for the lack of *myc* inhibition in my experiments is that the vivo-morpholinos (i) fail to enter sponge cells, or (ii) are trapped in endosomes after cellular uptake. To address this possibility, I suggest to use fluorescently-labelled morpholino oligos. Since vivo-morpholino oligos cannot be fluorescently labeled, this will imply to use regular morpholino oligos (less adequate for in-vivo settings as compared to the vivo-morpholinos used in this thesis). Also, I suggest shifting the morpholino target sequence in future experiments. Since morpholinos are typically very short, shifting their target sequence for a few base pairs can make a big difference in their efficiency.

Next, I transfected a Myc overexpression construct and a dn-Myc construct into sponge cells. This approach seemed promising since the technique to transfect and transcribe foreign DNA in *S. domuncula* cells was already established. Here, I managed to achieve even higher transfection efficiency of the constructs than previously shown[6,49] and both the newly made constructs can be used in future experiments. While I was able to detect *gfp* mRNA transcribed by the Actb::Gfp construct at similar levels as previously reported[6,49], I was not able to detect any transcription by the two newly generated constructs. The lack of detection of *gfp* mRNA transcription by these constructs may have several causes. For example, the *actb* intron, present in the Actb::Gfp construct but not in the Myc-overexpression or dn-Myc constructs, might include part of the transcriptional control. It has been shown that introns can have strong influence on the transcription efficiency of their locus[52]. An alternative explanation for the lack of detectable *gfp* mRNA expression by the two newly made constructs is that the presence of the *myc* sequence in the Myc-overexpression and dn-Myc constructs might interfere with the efficiency of transcription. Lastly, it could also be that Myc

overexpression and dn-Myc expression are toxic to the cells. Since only a small proportion of the cells of the explant are typically able to express Gfp from the exogenous DNA[6], the selective death of the cells overexpressing Myc or expressing dn-Myc would not have a noticeable effect on the morphology of the explants, or on the overall expression of *myc*. Finally, given that the level of expression of *gfp* mRNA by the Actb::Gfp construct is much lower than the level of *actb* mRNA expression[6], it is plausible that the majority of transfected DNA gets trapped in endosomes after uptake, and is therefore not accessible for transcription. One possible follow-up experiment to test this would be to transfect DNA in the presence of endosomolytic reagents.

Finally, I tried knocking down Myc using RNAi via shRNAs. This approach was interesting because its efficiency was previously shown in *Nematostella vectensis* embryos[37] using electroporation to deliver the shRNAs. I decided to use the transfection reagent INTERFERin instead of electroporation because of previous experiments which showed inefficient DNA transfection using electroporation[49]. However, I was not able to detect a reduction in neither *myc* nor *silicatein* mRNA levels. Several reasons could explain this result. Since only one set of conditions has been tried in my thesis, in the future I would suggest to use different amounts of shRNA and INTERFERin reagent.

Overall, going forward analyzing the function of Myc in sponge archaeocytes remains an interesting challenge. Once the necessary techniques are established (hopefully building up on the work presented here), it will be exciting to explore whether indeed, the co-expression of *myc* with some of the *S. domuncula* homologs of the human Myc direct targets represents a conserved Myc-regulated network that was important in the evolution of multicellularity.

References

1. Simion, P. *et al.* A Large and Consistent Phylogenomic Dataset Supports Sponges as the Sister Group to All Other Animals. *Curr. Biol.* **27**, 958–967 (2017).
2. Improved Phylogenomic Taxon Sampling Noticeably Affects Nonbilaterian Relationships. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2922619/.
3. Philippe, H. *et al.* Phylogenomics Revives Traditional Views on Deep Animal Relationships. *Curr. Biol.* **19**, 706–712 (2009).
4. Simpson, T. L. *The Cell Biology of Sponges*. (Springer New York, 1984).
5. Arnellos, A. & Keijzer, F. Bodily complexity: Integrated multicellular organizations for contraction-based motility. *Front. Physiol.* **10**, (2019).
6. Revilla-I-Domingo, R., Schmidt, C., Zifko, C. & Raible, F. Establishment of transgenesis in the demosponge suberites domuncula. *Genetics* **210**, 435–443 (2018).
7. Wilson, H. V. On some phenomena of coalescence and regeneration in sponges. *J. Exp. Zool.* **5**, 245–258 (1907).
8. Buscema, M., De Sutter, D. & Van de Vyver, G. Ultrastructural study of differentiation processes during aggregation of purified sponge archaeocytes. *Wilhelm Roux's Arch. Dev. Biol.* **188**, 45–53 (1980).
9. Borisenko, I. E., Adamska, M., Tokina, D. B. & Ereskovsky, A. V. Transdifferentiation is a driving force of regeneration in Halisarca dujardini (Demospongiae, Porifera). *PeerJ* **2015**, (2015).
10. Simpson, T. L. & Fell, P. E. Dormancy among the Porifera: Gemmule Formation and Germination in Fresh-Water and Marine Sponges. *Trans. Am. Microsc. Soc.* **93**, 544 (1974).
11. Funayama, N. The stem cell system in demosponges: Suggested involvement of two types of cells: Archeocytes (active stem cells) and choanocytes (food-entrapping flagellated cells). *Development Genes and Evolution* vol. 223 23–38 (2013).
12. Funayama, N., Nakatsukasa, M., Mohri, K., Masuda, Y. & Agata, K. Piwi expression in archeocytes and choanocytes in demosponges: insights into the stem cell system in demosponges. *Evol. Dev.* **12**, 275–287 (2010).
13. Müller, W. E. G. The stem cell concept in sponges (Porifera): Metazoan traits. *Seminars in Cell and Developmental Biology* vol. 17 481–491 (2006).
14. Borojevic, R. Étude expérimentale de la différenciation des cellules de l'éponge au cours de son développement. *Dev. Biol.* **14**, 130–153 (1966).
15. Musser, J. *et al.* Profiling cellular diversity in sponges informs animal cell type and nervous system evolution. *bioRxiv* 758276 (2019) doi:10.1101/758276.
16. Alié, A. *et al.* The ancestral gene repertoire of animal stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E7093–E7100 (2015).
17. Sebé-Pedrós, A. *et al.* Early metazoan cell type diversity and the evolution of multicellular gene regulation. *Nat. Ecol. Evol.* **2**, 1176–1188 (2018).
18. Sogabe, S. *et al.* Pluripotency and the origin of animal multicellularity. *Nature* **570**, 519–522 (2019).
19. Deng, W. & Lin, H. miwi, a murine homolog of piwi, encodes a cytoplasmic protein

essential for spermatogenesis. *Dev. Cell* **2**, 819–830 (2002).

20. Kuramochi-Miyagawa, S. *et al.* Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. *Development* **131**, 839–849 (2004).

21. Sánchez Alvarado, A., Newmark, P. A., Robb, S. M. C. & Juste, R. The Schmidtea mediterranea database as a molecular resource for studying platyhelminthes, stem cells and regeneration. *Development* **129**, 5659–5665 (2002).

22. Denker, E., Manuel, M., Leclère, L., Le Guyader, H. & Rabet, N. Ordered progression of nematogenesis from stem cells through differentiation stages in the tentacle bulb of Clytia hemisphaerica (Hydrozoa, Cnidaria). *Dev. Biol.* **315**, 99–113 (2008).

23. Reddien, P. W., Oviedo, N. J., Jennings, J. R., Jenkin, J. C. & Sánchez Alvarado, A. Developmental biology: SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science (80-.).* **310**, 1327–1330 (2005).

24. Ito, K. & Suda, T. Metabolic requirements for the maintenance of self-renewing stem cells. *Nature Reviews Molecular Cell Biology* vol. 15 243–256 (2014).

25. Bretones, G., Delgado, M. D. & León, J. Myc and cell cycle control. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* vol. 1849 506–516 (2015).

26. Iavarone, A. & Lasorella, A. Myc and differentiation: Going against the current. *EMBO Rep.* **15**, 324–325 (2014).

27. Dang, C. V. MYC, metabolism, cell growth, and tumorigenesis. *Cold Spring Harb. Perspect. Med.* **3**, (2013).

28. Hartl, M. *et al.* Stem cell-specific activation of an ancestral myc protooncogene with conserved basic functions in the early metazoan Hydra. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4051–4056 (2010).

29. Amati, B. *et al.* Oncogenic activity of the c-Myc protein requires dimerization with Max. *Cell* **72**, 233–245 (1993).

30. Nie, Z. *et al.* c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* **151**, 68–79 (2012).

31. Lin, C. Y. *et al.* Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* **151**, 56–67 (2012).

32. Muhar, M. *et al.* SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC axis. *Science (80-.).* **360**, 800–805 (2018).

33. Elkon, R. *et al.* Myc coordinates transcription and translation to enhance transformation and suppress invasiveness. *EMBO Rep.* **16**, 1723–1736 (2015).

34. Lascu, I. & Gonin, P. The catalytic mechanism of nucleoside diphosphate kinases. *Journal of Bioenergetics and Biomembranes* vol. 32 237–246 (2000).

35. Sebé-Pedrós, A. *et al.* The Dynamic Regulatory Genome of Capsaspora and the Origin of Animal Multicellularity. *Cell* **165**, 1224–1237 (2016).

36. El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432 (2019).

37. Karabulut, A., He, S., Chen, C. Y., McKinney, S. A. & Gibson, M. C. Electroporation of short hairpin RNAs for rapid and efficient gene knockdown in the starlet sea anemone, Nematostella vectensis. *Dev. Biol.* **448**, 7–15 (2019).

38. Takubo, K. *et al.* Regulation of glycolysis by Pdk functions as a metabolic checkpoint for cell cycle quiescence in hematopoietic stem cells. *Cell Stem Cell* **12**, 49–61 (2013).

39. Suda, T., Takubo, K. & Semenza, G. L. Metabolic regulation of hematopoietic stem cells in the hypoxic niche. *Cell Stem Cell* vol. 9 298–310 (2011).

40. Simsek, T. *et al.* The distinct metabolic profile of hematopoietic stem cells reflects their location in a hypoxic niche. *Cell Stem Cell* **7**, 380–390 (2010).

41. Kanatsu-Shinohara, M. *et al.* Myc/Mycn-mediated glycolysis enhances mouse spermatogonial stem cell self-renewal. *Genes Dev.* **30**, 2637–2648 (2016).

42. Oster, S. K., Mao, D. Y. L., Kennedy, J. & Penn, L. Z. Functional analysis of the n-terminal domain of the myc oncoprotein. *Oncogene* **22**, 1998–2010 (2003).

43. Mukherjee, S. *et al.* Japanese Encephalitis Virus-induced *let-7a/b* interacted with the <scp>NOTCH</scp> - <scp>TLR</scp> 7 pathway in microglia and facilitated neuronal death via caspase activation. *J. Neurochem.* **149**, 518–534 (2019).

44. Eisen, J. S. & Smith, J. C. Controlling morpholino experiments: Don't stop making antisense. *Development* vol. 135 1735–1743 (2008).

45. Wona, M., Ro, H. & Dawida, I. B. Lnx2 ubiquitin ligase is essential for exocrine cell differentiation in the early zebrafish pancreas. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 12426–12431 (2015).

46. Klann, M. & Seaver, E. C. Functional role of pax6 during eye and nervous system development in the annelid Capitella teleta. *Dev. Biol.* **456**, 86–103 (2019).

47. Lüscher, B. & Larsson, L. G. The basic region/helix-loop-helix/leucine zipper domain of Myc proto-oncoproteins: Function and regulation. *Oncogene* vol. 18 2955–2966 (1999).

48. Roulston, C. *et al.* '2A-Like' Signal Sequences Mediating Translational Recoding: A Novel Form of Dual Protein Targeting. *Traffic* **17**, 923–939 (2016).

49. Schmidt, C. MASTERARBEIT / MASTER ' S THESIS. (2018).

50. Yang, C., Qiu, L. & Xu, Z. Specific gene silencing using RNAi in cell culture. *Methods Mol. Biol.* **793**, 457–477 (2011).

51. Young, S. L. *et al.* Premetazoan ancestry of the Myc-Max network. *Mol. Biol. Evol.* **28**, 2961–2971 (2011).

52. Rose, A. B. Introns as Gene Regulators: A Brick on the Accelerator. *Front. Genet.* **9**, 672 (2019).

Figures