MQP-BIO-DSA-0204

Library MosSCI: A New Approach to Multiplexed Transgenic Strain Generation

A Major Qualifying Project Report

Submitted to the Faculty of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

in

Biology and Biotechnology

by

_____
Daniel Hartman

April 28, 2011

APPROVED:

_____
Sean Ryder, Ph.D.
Biochemistry and Molecular Pharmacology
UMASS Medical Center
Major Advisor

_____
David Adams, Ph.D.
Biology and Biotechnology
WPI Project Advisor

ABSTRACT

Early embryogenesis cell fate specification in *Caenorhabditis elegans* is driven primarily by post-transcriptional regulation, in which RNA-binding proteins bind to the 3' UTRs of complementary mRNAs. POS-1 is a RNA-binding protein necessary for germline specification in early embryogenesis. Transgenic worm strains with possible POS-1-binding sites were used to study the underlying mechanisms of progenitor cell fate regulation by POS-1. A previously developed Mos1-mediated single-copy insertion method was modified to increase the production efficiency of transgenic strains of *C. elegans* which will be used in future experiments.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank Sean Ryder for allowing me to work in his lab and supplying me with the necessary tools and support to successfully complete this project. I would also like to thank Brian Farley for teaching me the required techniques to carry out each experiment, as well as supplying me with plentiful background information and assistance during the course of my lab-work. Brian also provided me with the constructs inserted into the worms and performed each microinjection. I would also like to thank Brian and Sean for their aid in editing and finalizing both this report and my poster for the presentation. Finally, I would like to thank Prof. David Adams for his willingness to be my faculty advisor, assisting in the acquisition of this project, and providing me with timely feedback throughout the process.

BACKGROUND

Studies on nematodes, fruit flies, frogs, and mice have shown that oocytes undergo genetic reprogramming before fertilization, and that numerous cell fates are specified during early embryogenesis. However, cells at these stages of development are transcriptionally repressed. Therefore, from metazoans to mammals, post-transcriptional regulation of maternally supplied mRNAs is the driving force behind gameto- and embryogenesis. To further study the regulatory mechanics of early embryogenesis, the model *Caenorhabditis elegans* was chosen.

Caenorhabditis elegans

Caenorhabditis elegans are nematodes commonly used for genetic research for several reasons. The worms are easily maintainable in a lab setting due to their small size, short life cycle, simple storage, and large number of offspring per animal. Worms can be either male or hermaphroditic, more commonly the latter, which allows for either self or cross-fertilization (Riddle et al., 1997). The *C. elegans* germline is easily visible and made up of well-defined regions for each phase of reproduction, from gametogenesis through fertilization at the spermatheca. This visibility offers genetic researchers a model viable for fluorescence experiments (Farley and Ryder, 2008). Finally, the entire lineage of every cell in the body has been identified and is traceable from zygote to hatchling (Sulston et al., 1983).

Hermphroditic *C. elegans* produce both gametes in the same germline: spermatocytes are produced during the larval stage, while oocyte production begins following the transition into adulthood. Gametogenesis begins in the tip of the gonad furthest from the vulva (**Figure 1**) with mitosis of germ cell precursors.

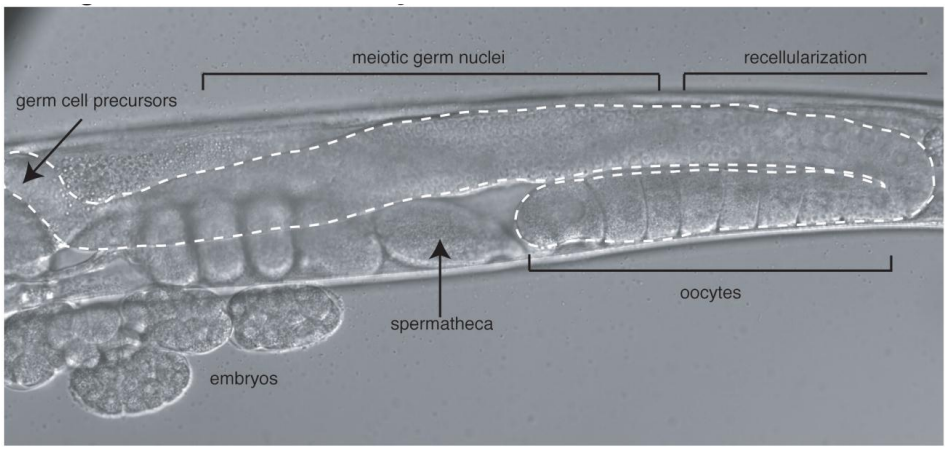

Figure 1: *C. elegans* Germline Anatomy (Farley and Ryder, 2008) (Picture is modified for germline outline).

After the cells divide, they pass into the meiotic phase where their membranes break down, forming a syncytium. The nuclei move to the walls of the gonad, encircling a cylinder of cytoplasm. Recellularization occurs at the germline "loop" where both spermatocytes and oocytes are formed at their respective stages (Farley and Ryder, 2008). The spermatheca is the last piece of the germline before the vulva and embryonic formation (see label in Figure 1) and stores sperm. The stored sperm can either be received from a male through copulation, from the producing hermaphrodite, or a mixture of both. In adulthood, gametogenesis switches from producing spermatocytes to creating oocytes, allowing fertilization to occur. As the oocyte passes through the spermatheca, it is fertilized by one of the two types of sperm present: produced or received. The resulting zygote quickly defines a body axis, separating anterior and posterior, in which the point of sperm entry determines the posterior end (Farley and Ryder, 2008).

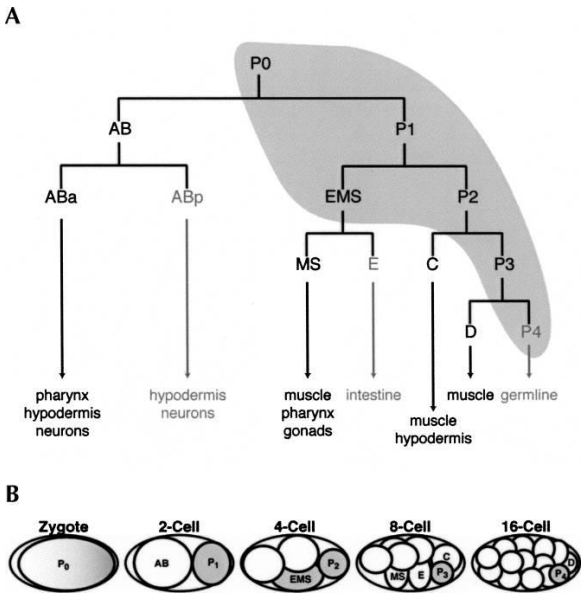

Figure 2: Early Embryogenesis Cell Fate (Farley et al., 2008)

Early embryogenesis (**Figure 2**) begins with the one-cell zygote (cell P0 in the figure) and ends at the 16-cell stage, when two germline cell descendants are formed (Panel B in the figure) (Tabara et al., 1999). For a majority of the cells in the early embryo, transcription is repressed. The zygote begins by dividing along the previously formed anterior-posterior axis, producing a larger anterior cell and a smaller posterior cell (2-cell stage). The anterior cell, AB, is a somatic founder cell and will divide once more to form two somatic blastomeres, ABa and ABp (Figure 2) (Farley and Ryder, 2008). At this point, the ABa and ABp cell fates have been specified, and transcriptional repression is turned off in the anterior cells.

The posterior cell (P1) is the progenitor of the germline and follows a stem cell-like lineage, with each division yielding an anterior founder cell and a posterior germline blastomere. The final division of the posterior cell occurs at the 16-cell stage, and produces two germ cell descendants, Z2 and Z3, ending early embryogenesis (P_0-P_4, Figure 2) (Tabara et al., 1999). Transcription is repressed in the posterior cells throughout the entirety of this process. Thus, cell

fate specification is driven by post-transcriptional regulation of maternally-supplied mRNAs (Farley and Ryder, 2008).

Post-transcriptional Regulation

Maternal mRNAs are mRNAs produced by the maternal genome and packaged into oocytes for use in embryogenesis. Germline cell fate specification is known to take place in the posterior blastomeres during early embryogenesis, though, the cells are transcriptionally silenced. The lack of transcription means that the only mRNAs to be acted upon are those supplied before fertilization, the maternal mRNAs. Therefore, cell specification is controlled by post-transcriptional regulation of the maternal mRNAs (Farley and Ryder, 2008). Post-transcriptional regulation, more specifically translational regulation, is a process where RNA-binding proteins interact with mRNA to control translation of proteins or affect mRNA stability. Research shows that the targets for RNA-binding proteins in the germline progenitors are the mRNA 3' UTRs, possibly due to the role 3' UTRs play in mRNA stability and the lack of ribosomal interference (Merritt et al., 2008). The 3' UTR lies outside the coding region of the mRNA, allowing proteins to bind without interfering with the large ribosome complex, which could strip the protein from the mRNA. Also, 3' UTRs aid in capping the coded sequence with a poly-A tail, improving mRNA and protein stability. While the full mechanism of regulation is not completely understood, evidence shows that the vast majority of identified regulatory events in *C. elegans* embryogenesis are inhibitory. Studies also noted that proper expression of MEX-3 and GLD-1, two important RNA-binding proteins, was dependent on the presence of target 3' UTRs and that expression was independent of promoter specificity (Merritt et al., 2008).

Many maternal mRNAs and RNA-binding proteins have been identified, as well as possible binding sites, consensus sequences, and the effects of lacking one or the other. For example, *glp-1* and *apx*-1 are genes that have been found to be translationally regulated during embryogenesis (Ogura et al., 2003). MEX-3, GLD-1, and POS-1 are three RNA-binding proteins whose consensus sequences have been identified, which has allowed for the identification of numerous possible target 3' UTRs (Carmel et al., 2010; Farley et al., 2008; Pagano et al., 2009). More specifically, POS-1 has been shown to interact with *glp-1* at its spatial control region of the 3' UTR, with no binding occurring if the region is mutated. This interaction prevents the translation of *glp-1* in embryonic posterior cells. However, this is only one example of thousands of possibilities. Studies of POS-1 have revealed that it regulates many more maternal mRNAs than *glp-1*, and plays a major role in germline cell fate.

POS-1

POS-1 is a cytoplasmic RNA-binding protein containing two copies of the CCCH zinc finger motif. POS-1 uses the CCCH finger motifs to bind to target 3' UTRs containing a sequence-specific site (highest affinity UA(U_{2-3})RD(N_{1-3})G) (Farley et al., 2008). Mutants lacking POS-1 are maternal-effect embryonic lethal, have ectopic pharynx, no intestine, and no germline precursors. POS-1 mutants also have defects in germline blastomeres, in which cells experience abnormal cleavage and P granule distribution, and shorter cell cycle. Without POS-1, posterior cells fail to divide into separate germline and somatic cells, resulting in no germ cell precursors (Tabara et al., 1999). During the first cleavage of the embryo, POS-1 is asymmetrically divided, with a higher concentration in the posterior daughter cell; each

subsequent division undergoes the same effect. Therefore, POS-1 is required for germline cell fate specification.

The POS-1 consensus sequence has been identified, but the mRNAs that it directly regulates are still unknown. POS-1 is known to negatively regulate *glp-1* expression in embryonic cells by binding to two separate regions (**Figure 3**) in the *glp-1* 3' UTR, and that binding to the region requires an intact second finger motif (Ogura et al., 2003). There is also a possibility that POS-1 forms complexes with other RNA-binding proteins, such as GLD-1, to increase regulatory efficiency (Farley et al., 2008).

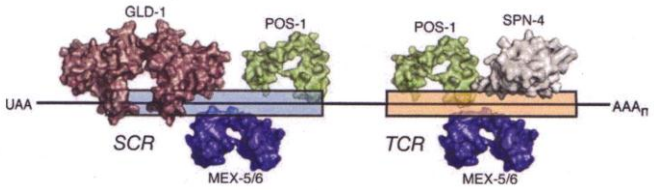

Figure 3: Layout of *glp-1* regulatory proteins (Farley and Ryder, 2008)

These are two examples of POS-1 regulation, though the mechanics behind how POS-1 interacts with each mRNA and protein are not understood. However, POS-1 binding sites are very common and are not only found in thousands of different 3' UTRs, but that the majority of those 3' UTRs contain multiple POS-1 sites within close proximity (**Figure 4**). This means that multiple POS-1 proteins could be needed to regulate maternal mRNAs and drive germline cell fate (Farley et al., 2008).

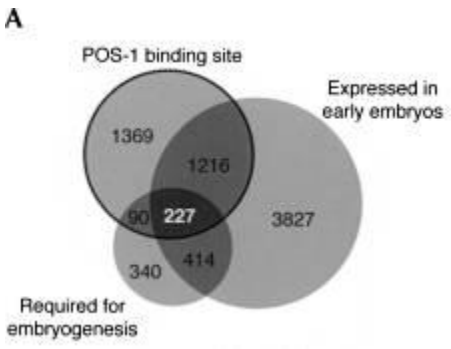

Figure 4: Venn diagram representing 3' UTRs possibly regulated by POS-1 (Farley et al., 2008)

Experiments designed to identify the key factors of POS-1 regulation in early embryogenesis use transgenic strains of *C. elegans* containing target 3' UTRs and GFP markers. Non-transgenic worms are difficult to perform translational regulation experiments due to the need for immunofluorescent staining, added reagents, and the inability to effectively mutate the 3' UTR. By inserting a transgene containing GFP, germline and embryo expression patterns can be created and compared to the patterns seen in POS-1. To determine if POS-1 and the inserted 3' UTR interact, though, either *in vivo* or *in vitro* experiments can be performed. *In vivo*¸ POS-1 is knocked-out in the transgenic strain, and the new expression pattern identified. This pattern can then be compared to a pattern received by mutating the POS-1 binding site(s) of the 3' UTR, and if the patterns are the same or comparable, then the conclusion can be drawn that POS-1 and the maternal mRNA 3' UTR interact. The *in vitro* experiment is comprised of running gel-shift assays. The gel-shift assay tests for binding affinity of POS-1 to target 3' UTRs. First, the wild-type 3' UTR is run with POS-1, followed by an assay using a 3' UTR with mutated POS-1 binding sites. The affinities of both assays are compared, and if the affinity of the wild-type is significantly higher than the mutated 3' UTR, then POS-1 binds to that maternal mRNA. The

downside to these experiments, however, is that thousands of transgenic strains would be required to perform these experiments on the multitude of possible POS-1, maternal mRNA combinations.

Mos1-mediated Single-copy Insertion

Currently, there are two commonly used forms of creating transgenic strains of *C. elegans*: bombardment (biolistic transformation) or microinjection. Microinjection uses needles to inject the plasmid of interest directly into either the oocytes or meiotic syncytium. Microinjection has the advantage of being cheaper and requiring less time overall, making it the more frequently used method (Rieckher et al., 2009). There are some significant disadvantages to using microinjection, however. The DNA injected into a worm exists in the form of extrachromosomal arrays which are not integrated into the worm chromosome, and instead form a "minichromosome". These minichromosomes are unstable in meiosis and mitosis, often leading to mosaic offspring, where not each cell expresses the transgene. Also, each array contains numerous copies of DNA leading to possible over-expression or toxic effects. The high copy number also causes difficulties in germline studies due to the *C. elegans*' natural ability to silence repetitive arrays in the germline (Frøkjaer-Jensen et al., 2008).

One way to avoid silencing and mosaic worms following microinjection is to integrate the DNA into the worm's chromosome. Integration is promoted by causing a double-stranded break in the chromosomal DNA, forcing the chromosome to repair itself with nearby DNA, which could be the injected plasmids or its own DNA. Radiation or mutagenic chemicals can be used to cause the double-stranded break, but the system is not full-proof (Rieckher et al., 2009).

Silencing can still occur in integrated transgenes, possibly due to multiple copies of DNA inserted at different points of breakage or transcriptional silencing (Frøkjaer-Jensen et al., 2008).

The second method of transgenic strain generation is capable of chromosomal integration without the need for radiation or chemicals. Bombardment is the process of coating gold microparticles with the DNA of interest and then "shooting" them into the worm at high speeds by way of a "biolistic bombardment device". The advantages of bombardment include low DNA copy number, ease of performance, and possible chromosomal integration. However, bombardment is the less frequently used of the two methods due to its expense, time consumption, and chance of causing delocalized expression (Rieckher et al., 2009).

There is a common disadvantage of both of these procedures, which makes their use in POS-1 experiments ineffective: random integration. While radiation and chemicals can integrate arrays into the worm genome, the number and location of the integrations is uncontrollable. Radiation and chemicals cannot be direct to specific locations of the chromosome, possibly causing drift expressions and multiple copies of integration. This makes mutating target 3' UTRs difficult and unreliable. Bombardment, whether integrated by itself or through radiation/chemicals, suffers from the same random integration (Rieckher et al., 2009).

To avoid the fallouts of microinjection and the expense of bombardment, a new procedure was designed (Frøkjaer-Jensen et al., 2008) that creates transgenic strains of *C. elegans* with a single-copy of integrated DNA. The method, called Mos1-mediated single-copy insertion (MosSCI), uses microinjection to deliver the arrays of transgenes to the worm along with several other components to allow for selection and integration. Integration in MosSCI requires the presence of one copy of the *Drosophila* Mos1 element in the *C. elegans* genome, and avoids the hazards of radiation and mutagenic chemicals (Frøkjaer-Jensen et al., 2008).

Mos1 is a transposon that can be inserted at defined loci of the *C. elegans* chromosome, and then excised by a Mos1 transposase expressed from a heat-shock promoter (Bessereau et al., 2001) Mos1 sites are chosen by their neutrality: no interference on adjacent gene function, and no presence of enhancers or promoters that would affect transgene expression. The most commonly used site is found on chromosome II, in a tail-to-tail orientation at the *ttTi5605* Mos1 allele. It is important to prepare a transgene that is viable for MosSCI, generally 7kb or less in length, and that will be suitable for the intended experiments (Frøkjaer-Jensen et al., 2008).

For POS-1 studies, a transgene was designed that contained three parts: a *mex-5* promoter, a component of MODC PEST::GFP::histone2b, and the target 3' UTR. The *mex-5* promoter is used due to its presence throughout the germline, which is necessary for creating germline expression. The GFP is targeted at histone2b to ensure that the expression is restricted to the nucleus. The addition of MODC PEST (mouse ornithine decarboxylase) is designed to degrade protein to stop GFP drift out of the nucleus and into other areas of the germline. The 3' UTR is then chosen based off of whether or not it contains target binding sequences for POS-1. The entire mRNA does not need to be integrated due to translational regulation occurring only by RNA-binding protein interactions with the 3' UTR.

The entire array injected into each worm contains positive-selection markers, negative-selection markers, the target transgene, and Mos1 transposase expressed by a heat-shock promoter (**Figure 5**). The *C. elegans* used for injection are *unc119(-)* which causes poor coordination, manifested as non-wild-type movement. To recover coordination, *unc119(+)* is attached to the target transgene. This complex is bordered by a left and right homology arm, which contains ~1.4kb of DNA homologous to genomic DNA adjacent to the Mos1 transposon; this entire construct is the intended integrant at completion. The extrachromosomal array,

intended to be lost following heat-shock, contains *twk-18i* (causes paralysis at 25°C) and mCherry (targeted for the pharynx and body wall). These two components act as negative and positive selections at different time points throughout the procedure. The last part of the array is Mos1 transposase expressed from a heat-shock promoter, allowing for excision of Mos1, causing a double-stranded DNA break and the uptake of the target transgene (Frøkjaer-Jensen et al., 2008).

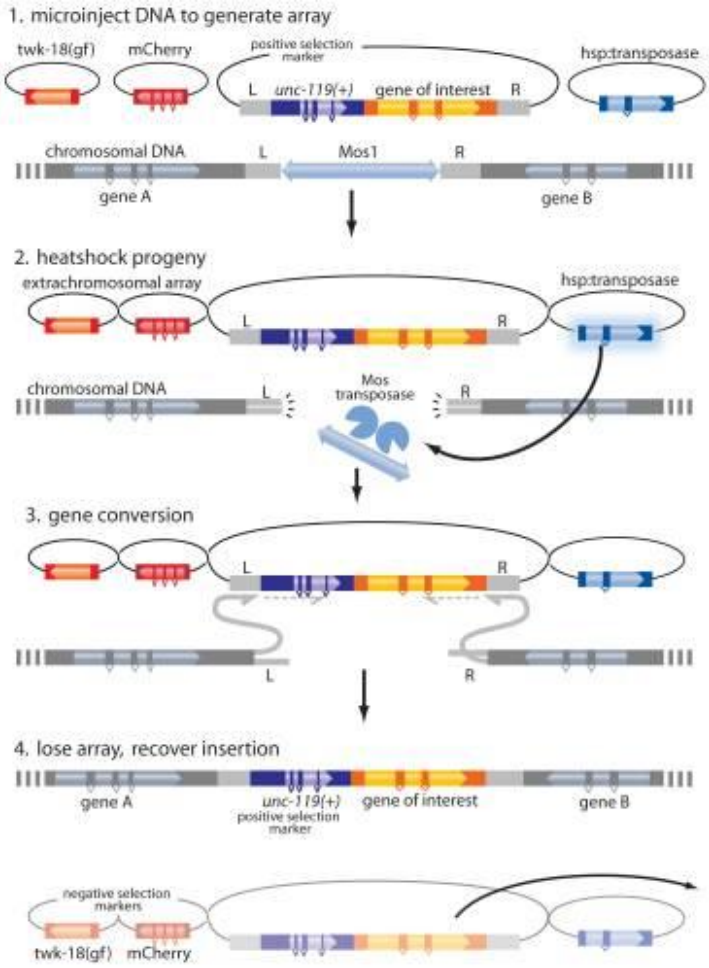

Figure 5: Internal Process of MosSCI (Frøkjaer-Jensen et al., 2008)

MosSCI has been shown to work most effectively with transgenes of size 7kb or smaller, at an efficiency rate of one successful insertion for every twenty worms injected. The majority of these insertions are single-copies at the defined Mos1 location. Most importantly, strong germline expression can be achieved through MosSCI with limited to no silencing over several generations; sperm expression is weak and almost exclusively seen in hermaphrodites (Frøkjaer-Jensen et al., 2008). However, this process takes anywhere from 2-4 weeks to complete and yields only one transgenic *C. elegans* strain. To fully study every combination of POS-1 and possible target 3' UTRs, thousands of transgenic strains would be required. The time requirement to achieve this would be too great to be possible. The solution to this is to modify MosSCI to produce multiple strains of transgenic worms at one time.

The approach taken to modify MosSCI has been named Library MosSCI. The idea behind Library MosSCI is that instead of inserting an array containing only one transgene of interest, a "cocktail" of transgenes is inserted. For this project, nine maternal mRNAs were selected based on their possession of multiple binding sites containing the POS-1 consensus sequence. These binding sites were all found in the 3' UTR and were within 20 nucleotides of each other. The intended result is multiple transgenic worms containing single-copies of different 3' UTRs that can be used in future POS-1 translational regulation experiments.

PROJECT PURPOSE

The purpose of this project was to modify the Mos1-mediated single-copy insertion method of Frøkjaer-Jensen et al. (2008) to produce multiple, different strains of transgenic *C. elegans* in the same timeframe as creating one strain using the original procedure. This modification would allow for quicker production of a wider range of transgenic worm strains useful in studying the mechanics behind translational regulation of maternal mRNAs in germline cell fate specification. The second project goal was to generate transgenic worm strains containing 3' UTRs with possible POS-1-binding sites. These strains would contain green fluorescent protein alongside the target 3' UTR, which could be used in further exploration of how POS-1 interacts with mRNA in various cells of the germline.

METHODOLOGY

Plasmid Transformation

To produce sufficient plasmid for the injection library, nine 3' UTR constructs (*set-21, hbl-1, cwn-1, set-6, usp-14, atg-4.2, kin-25, mex-3, cul-4*) and the four extrachromosomal array constructs [pCFJ70 (*twk-18*), pCFJ90 (pharynx mCherry), pCFJ104 (body mCherry), and pJL44 (Phsp::transposase)] were transformed into *E. coli* and cultured overnight.

Plasmid Purification and Ethanol Precipitation

The plasmids previously transformed into the bacterial cultures were purified using the QIAprep Miniprep Kit. The purified plasmids were then further purified by ethanol precipitation using 5M NaCl and 200-proof ethanol.

Microinjection Preparation

Plasmid concentration and library creation

The concentrations (ng/µl) of all thirteen purified plasmids were calculated, a library containing 1µg of each plasmid was constructed, and then the final concentration of the library was determined.

Injection mix

Into a single microcentrifuge tube was added 50ng/µl of both pJL44 and the previously constructed library. To the same tube, 10 ng/µl of pCFJ70, 5 ng/µl of pCFJ104, and 2.5 ng/µl of pCFJ90, were added. The final volume of the mixture was then brought to 20 µl using 2 µl of 1X

Buffer (2% polyethylene glycol, 8000 molecular weight, 20 mM potassium phosphate pH 7.5, 3 mM potassium citrate, pH 7.5) and the remaining difference of volume with filtered water. The mixture was then injected into ~30 worms.

Mos1-Mediated Single-Copy Insertion

Pre-heat-shock: Propagation and Screening

After the worms were injected with the array, they were left to propagate at 25°C on 60mm RNAi plates for two days, at which point they were checked for starvation; at this point, a set of five backup plates were set aside containing pre-heat-shock worms in case of emergency. RNAi plates contained IPTG and Ampicillin, and were lined with *twk-18* RNAi positive *E. coli*. The worm populations were checked regularly every two days to ensure starvation did not occur, as it could lessen the chances of insertion. After four days of propagation at 25°C, worms were screened for the presence of mCherry using a fluorescence microscope. Six plates contained populations of primarily mCherry positive worms and were cut into quarters, each quarter was moved to a new RNAi plate for further propagation. Eight days later, approximately three generations of progeny (twelve days) from the original injected worms, plates that had more than 30, mCherry positive young adults were heat-shocked.

Heat-shock

15 plates contained thirty or more mCherry positive young adult worms. The plates were wrapped in parafilm until completely sealed, then submerged in a 34°C water bath for one hour. After the one hour of heat-shock, the worms were moved to a 15°C incubator to recover for two hours. Worms were then removed from the plates by washing with approximately 2 mL of

filtered water and storing the wash in a 15 mL centrifuge tube. After all of the worms were removed from each plate, the majority of the worms were allowed to settle to the bottom of the tube ensure they were clean of RNAi bacteria food. The water was then aspirated off, and another 10 mL of filtered water added. This wash cycle was performed two more times. Following the final aspiration, worms were pipetted onto new non-RNAi (normal) 60 mm plates at a volume that yielded ~20 worms per plate; 24 plates were obtained.

Post-heat-shock: Propagation and Screening

Heat-shocked worms were propagated at 25°C (to activate any possible remaining *twk-18* for negative selection) for two days, and then expanded from 24 plates to 48 plates by using two quarters of each original plate. The expanded 48 plates were then propagated for three more days at 25°C, at which point the plates were screened for mCherry-negative, wild-type moving adults. 21 non-red, wild-type worms from eight different plates were found and each picked onto their own 35mm plate. Plates were labeled A-H for each plate picked from (putative strains), then divided into numbers for individual worms. These plates were then allowed to propagate indefinitely (moving to new plates of food as needed) while PCRs were performed.

Polymerase Chain Reactions

Worm lysis

For each round of PCR, DNA was extracted from the worms by using a thermal cycler. To do this, two adult worms were picked into 5 µl of a 1:20 mixture of proteinase K and 30 mM Tris pH 8.8 for each reaction. The worms were then frozen on dry ice for approximately ten

minutes. The thermal cycler was set to run for one hour at 65°C to allow digestion by proteinase K, followed by fifteen minutes of 95°C to denature the proteinase K. Then the worms were removed from the dry ice, place in the cycler and the program was run. The resulting solutions could then be used in the desired PCR reactions.

Shorter amplification primer pair – Taq

PCR was used on all 21 strains of worms received after heat-shock, as well as the pre-heat-shock worms, and all were run against a positive strain which contained a known integrant and lacked an extrachromosomal array. For each reaction, there was 5µL each of DNA from lysed worms, 0.2mM dNTP, 0.5µM primer 1 (BMF 69, halfway through GFP), 0.5µM primer 2 (BMF 479, partway through right homology arm), and ThermoPol buffer; there was also 1µL of Taq polymerase and 24µL of filtered water. The mix and extractions were kept on ice separately until the thermal cycler was pre-heated for the short-taq program, at which point 45µL of the master mix was added to each tube of extraction, the tubes placed in the thermal cycler, and the sequence started.

3' UTR sequencing and genomic DNA extraction

The strains that received positive results from the shorter amplification PCR were sequenced and ran through the BLAST system to determine which 3' UTRs were present in the putative strains. The four different 3' UTR strains obtained were separated onto their own plates and propagated for five days. After checking for at least 20 adult worms per plate, the four plates were washed with TE, and the worm-wash mixtures moved to microcentrifuge tubes. The tubes were centrifuged, the supernatant removed, then the pellet was frozen. A 1mL mixture of

200µg/ml Proteinase K in 2x NTE (200mM NaCl, 100mM Tris(pH8.5), 40mM EDTA) was prepared. 200µL each of TE and the previous solution were added to the thawed pellets, which were then incubated at 65°C for three hours, with gentle agitation every 30 minutes. Extraction was performed with two washes of PCI and one wash of chloroform, followed by the addition 20µL 5M NaCl and 400µL 100% iso-propanol to each tube. The solution was mixed and ethanol precipitated. Finally the pellet was resuspended in 100µL 0.1x TE and the concentrations of each DNA solution calculated using a spectrophotometer.

Longer amplification primer pair – Taq and Elongase

A primer pair that amplified a longer product was used on the previously extracted genomic DNA with both Taq polymerase reactions and Elongase reactions. The Taq polymerase PCR was performed following the same protocol as the shorter primer pair. The Elongase protocol required the master mix be kept cold and split up until addition to the 5µL of genomic DNA. In one tube, 5µL each of 0.2mM dNTP, 0.2µM primer 1 (BMF69), 0.2µM primer 2 (BMF 480, located in the genomic DNA outside the right homology arm), and filtered water was added. The second contained 1µL of Elongase, 10µL of Buffer B (300 mM Tris-SO_4, (pH 9.1 at 25°C), 90 mM $(NH_4)_2SO_4$ and 10 mM $MgSO_4$), and 19µL filtered water. The two master mixes were then combined and 45µL added to each tube of 5µL of genomic DNA, and placed in the thermal cycler for the "Elongase" program.

RESULTS

The goal of this project was to modify the Mos1-mediated single-copy insertion method to produce more than one type of transgenic strain of *C. elegans* in the same amount of time it takes to produce one strain using the original MosSCI method. To do this, a library of 3' UTRs containing multiple possible POS-1 binding sites was used. The purpose behind using POS-1 targets was to end up with transgenic strains that could be used in future experiments to explore the mechanics of POS-1 translational regulation during early embryogenesis.

The selection of worms for propagation both before and after heat-shock was key in conserving time and materials while maximizing the efficiency of the overall method. Following microinjection and propagation of the injected worms over three generations, the worms were screened for the presence of mCherry (pharyngeal and body fluorescence) and wild-type movement (sine-wave shape) (**Figure 6**). The goal was to have a majority of adult worms showing these rescued attributes, as it meant that the array was present and passed on to progeny. Worms without the array could be easily identified by their lack of red fluorescence and C-shaped bodies (note the worm on left in Figure 6). Out of 30 injected worms, 24 worms showed rescued lines, and 18 of those lines stably propagated the array.

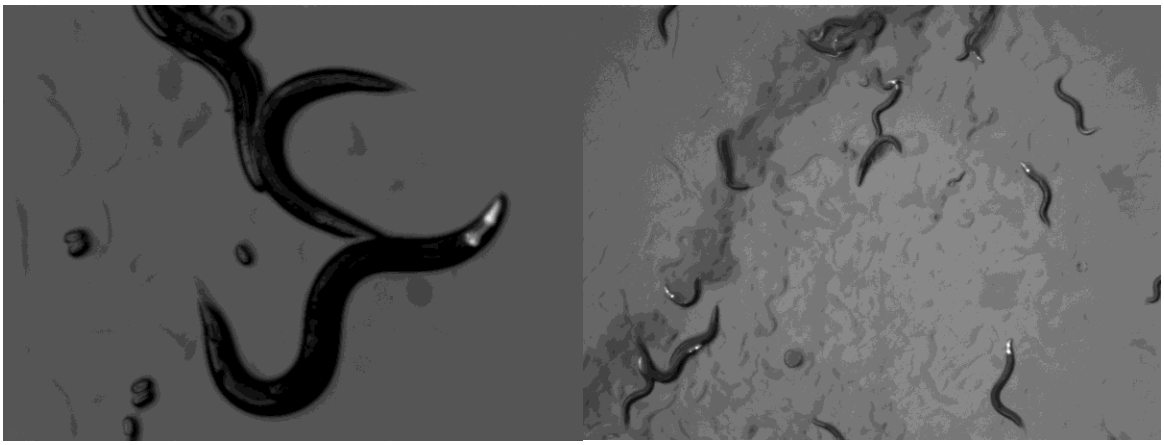

Figure 6: Photos of array-positive and array-negative worms pre-heat-shock.

To obtain different transgenic strains, it was important to have at least two or more 3' UTRs present in the array before heat-shock; all nine would be preferable. A PCR using Taq polymerase and the primer pair BMF69-479, which amplifies a region containing half of the GFP marker, the 3' UTR, and part of the right homology arm, showed that at least seven of the nine 3' UTRs were present in the worms pre-heat-shock (**Figure 7**) denoted by multiple similar-sized bands; a 2-Log DNA Ladder was used for all gels. The positive lane contains a strain of *C. elegans* with a known integrant and lacking the extrachromosomal array, while the sample lanes are a concentration gradient of the number of worms lysed, ranging from one through five worms, left to right.

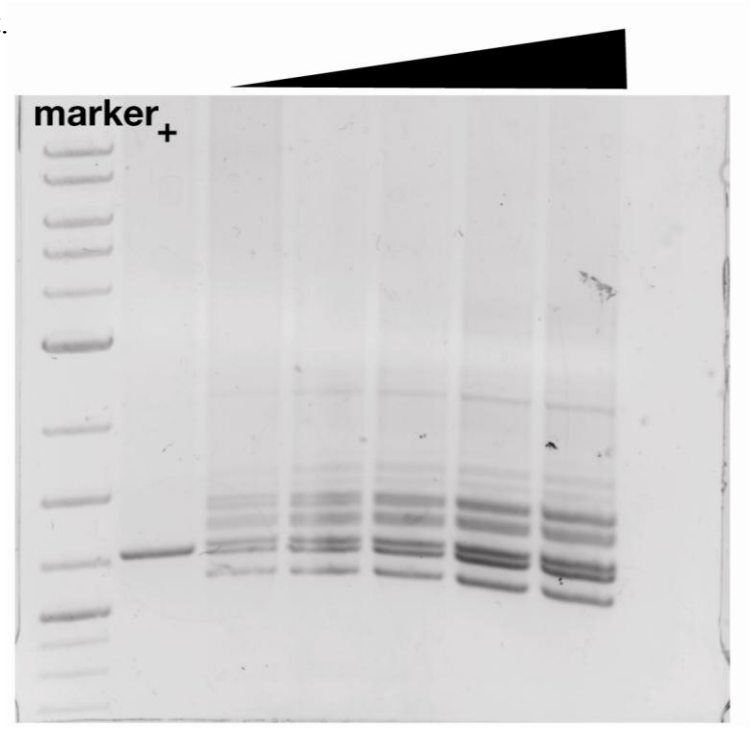

Figure 7: PCR of worms pre-heat-shock to show presence of multiple 3' UTRs.

15 plates were heat-shocked, each containing at least 30 array-containing young adult worms. After heat-shock and wash, 24 plates were obtained with ~20 adults each. Screening of heat-shocked plates yielded 21 non-red, wild-type young adults out of 480. The 21 worms were

picked from 8 different plates. Worms of the same plate were assumed likely to contain the same 3' UTR; however, each of the 21 strains was studied individually.

Another PCR, the same as used in Figure 7, was performed on 20 of the 21 strains (one strain failed to propagate) to test for the presence of 3' UTRs in the worms, but not necessarily for chromosomal integration. The primer pair BMF69-479 amplifies a region that does not contain the worm's genomic DNA, and therefore could detect a transgene that still existed as an array. The other purpose of this PCR is to ensure that only one 3' UTR from the entire library is present. The results can be seen in **Figure 8**, showing multiple lanes with single bands of different sizes. These results suggested that not only were different 3' UTRs retrieved, but that each worm only contained one 3' UTR from the original library.

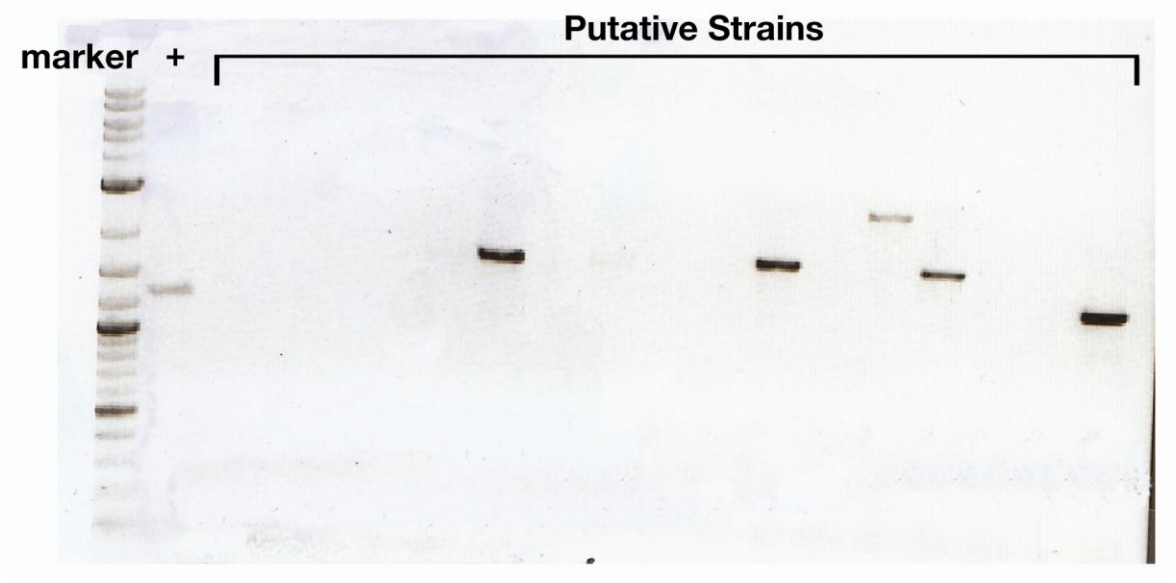

Figure 8: PCR of putative strains post-heat-shock to check for presence of a single 3' UTR.

The PCR products from each positive strain were sent for DNA sequencing, so that each strain could be attached identified. Four out of the nine injected 3' UTRs were retrieved from the sequenced DNA: mex-3, cwn-1, hbl-1, and kin-25. The returned sequences were then compared

to the accepted sequences for each 3' UTR to determine the level of consensus. Possible mutations were noted in three of the four 3' UTRs, but kin-25 showed a complete consensus, as seen in **Figure 9**. It is possible, however, that the original 3' UTRs are not mutated, but appear so due to the sequencing of PCR products, in which the amplification process could have input errors. Sequencing using less processed DNA could show more aligned sequences and less "mutations". With the 3' UTRs sequenced, the next step was to see which were integrated into the worm chromosome.

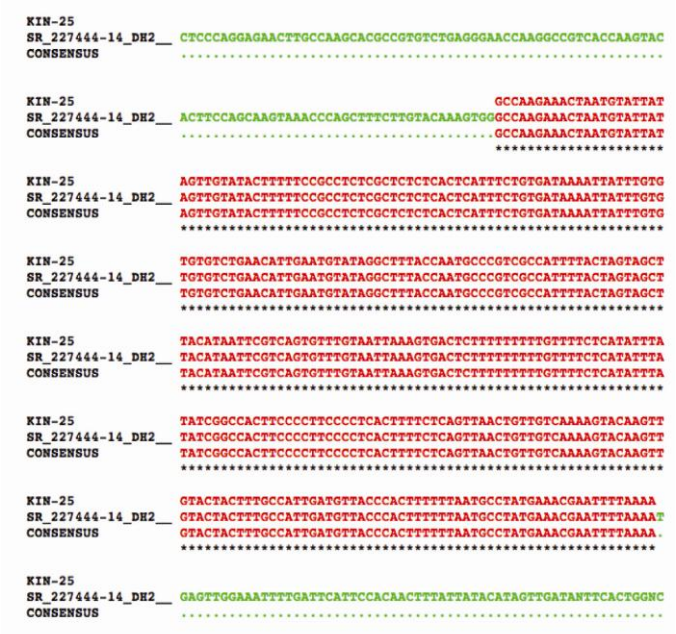

Figure 9: Sequence Comparison of Known and Obtained *kin-25* 3' UTR.

Genomic DNA extractions were performed on the worm strains containing each 3' UTR. The genomic DNA was then analyzed by PCR, one with both Taq polymerase (data not shown) and one with Elongase (**Figure 10**). Both PCR's used the primer pair BMF69-480 which amplifies a region containing half of the GFP marker, the entire 3' UTR and right homology arm, as well as a section of the worm genome outside the homology arm. This region of amplification proves integration by including worm genomic DNA. If the transgene were not integrated, the

upstream primer (in the GFP) would amplify without the downstream primer, causing smears in the lanes. However, integration would cause the downstream primer (located in the genomic DNA) to control amplification, resulting in a single band. Figure 10 shows a lane for each of the sequenced 3' UTRs along with a positive of known integration. Each of the lanes contains only one band, each different in size from the others. This means that four of the nine 3' UTRs from the library successfully integrated into the worm chromosome. Library MosSCI obtained 4 transgenic worm strains for 30 injected worms (13.3%) compared to the 1 worm for every 20 injected (5%) with MosSCI. In addition, the 4 transgenic strains contained four different transgenes, increasing efficiency even further.

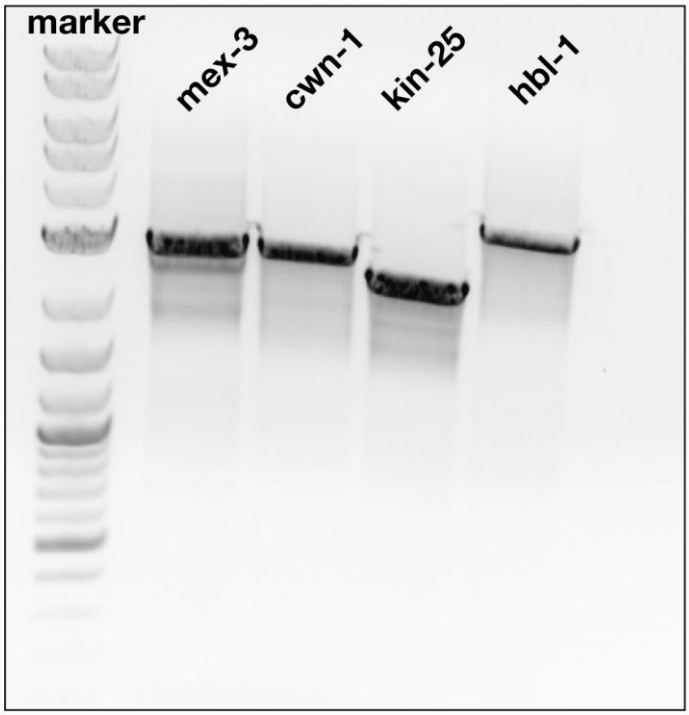

Figure 10: PCR of positive strains to check for integration of transgene into worm genome.

To determine if only one copy of transgene was integrated, the PCR product lengths from Figure 10 were compared to the known lengths for the amplified region. The only varying lengths were the 3' UTR lengths, while the homology arm (1.5kb), GFP (500kb), and genomic

DNA (<50kb) lengths were the same for each transgene. In general, the 3' UTRs ranged between 100 and 1000 nucleotides, creating products in the 2000-3000kb range. **Table 1** contains the PCR length and the expected length for each 3' UTR. Based off of these results, each transgenic strain contained only a single copy of the corresponding 3' UTR.

3' UTR	PCR Length (~kb)	Expected Length (entire transgene) (kb)
kin-25	1500	2320
cwn-1	2000	2182
hbl-1	2500	3148
mex-3	2000	2483

Table 1: Comparison of known transgene amplification lengths to the obtained PCR lengths.

The final goal of this project was to obtain worms with different integrated transgenes that expressed nuclear GFP in the germline. However, screening of ~20 worms from each positive strain in Figure 8 showed no fluorescence. The worms were propagated for ~10 generations, in an attempt to reverse possible silencing, with screening of at least 20 worms at the 5 and 10 generation mark, both with no nuclear GFP expression. One worm showed oocyte cytoplasmic GFP, but this was not reproduced in other worms and was concluded to be irrelevant.

DISCUSSION

Multiple, varying transgenic strains of *C. elegans* containing a single-copy of a 3' UTR at a defined locus of the chromosome were obtained through a modified Mos1-mediated single-copy insertion method. This modification occurs at the point of preparing the injection mixture, where nine transgenes were combined instead of one. The newly named Library MosSCI yielded four different transgenic strains (cwn-1, kin-25, hbl-1, mex-3) out of the nine 3' UTRs injected, in the same time the original MosSCI method would generate one strain. The strains contained a single-copy of the 3' UTR fully integrated into the chromosome; however, no germline GFP expression was seen. The secondary goal of this project was to generate transgenic strains containing 3' UTRs with multiple, possible POS-1 binding sites that could be used in future translation regulation studies, but the experiments require GFP expression.

The original MosSCI method was able to gain germline expression following the same procedure as that used in Library MosSCI, with the exception of additional transgenes (Frøkjaer-Jensen et al., 2008). *C. elegans* are able to silence trangenes in their germline by way of different types of small RNA (siRNA) pathways (Zhang et al., 2011). In some cases, the silencing mechanisms recognize repetitive sequences of DNA. Multiple components are involved in the siRNA process, including piRNAs (silence transposons), RNA-dependent RNA polymerases (RdRPs), Dicer, and worm-specific Argonaute proteins (WAGOs) (Gu et al., 2009). One key element in a prominent germline silencing pathway is the protein Dicer-related helicase 3 (DRH-3). DRH-3 is involved in the biogenesis of an RdRP called 22G-RNA, which targets exogenous genes in the germline. Mutations in DRH-3 helicase motif have shown knock-outs of RdRPs as well as other siRNAs, reducing the silencing effect. However, the downside to this

mutation is the loss of 22G-RNA, which is involved in euchromatic chromosome segregation (Matrange and Ryle, 2010).

For Library MosSCI to achieve an improved outcome over the original MosSCI, GFP germline expression is required. Several options are available to regain or promote expression. The first option is propagating the transgenic strains for multiple generations. This allows the siRNAs to dilute out of the system as less and less become transcribed and packaged into oocytes. However, after ~10 generations of the transgenic strains obtained in this project, no reversal of silencing was achieved. The second option is to change the injection mix to have a lower concentration of DNA for each plasmid, which would lessen the repetitive sequences recognized by siRNAs. Along those lines, additional DNA of various types could be added to the injection to dilute the repetitive sequences, with the same effect. The last option would be to continue experiments on mutations in the DRH-3 protein or other various RdRPs, such as 22G-RNA. Mutations could occur in the recognitions sites of 22G-RNA for the silencing pathway, allowing for normal chromosomal segregation function while eliminating the silence effect. No matter the method used, it would be beneficial to optimize Library MosSCI to circumvent germline transgene silencing.

Library MosSCI can still undergo further experimentation and optimization to obtain not only stronger germline GFP expression, but also generate more strains per injection. Careful monitoring of worms during propagation and selection may increase chances of proper insertion. The final working library seen here was the third trial of modifying MosSCI, but the important difference between the first two trials and the final one was worm propagation. The original method for selection and propagation was worm bleaching, which was later determined to select for worms not containing the pre-heat-shock array. Worm bleaching consists of washing worms

off of plates, destroying the live worms with bleach and sodium hydroxide, and separating out the remaining embryos. However, the worm embryos that were more likely to survive were embryos still within the parent, parents who lack the ability to release their embryos due to paralysis. Switching over to a method that selected for non-paralyzed worms (quartering onto new plates) proved to increase the chances of insertion. Library MosSCI may also be capable of undergoing expansion to allow insertion of hundreds of transgenes per worm by lowering plasmid concentrations in the injection mixture. Lastly, it is possible that DNA repair favors shorter transgenes due to *cwn-1*'s higher frequency of insertion and shorter length compared to the other three 3' UTRs. If transgene size plays a role in insertion frequency, then injection mixes can be rearranged to favor one transgene over another.

Library MosSCI offers the chance to generate hundreds of transgenic strains per round of microinjections. In this case, thirty worms were injected and four different strains were obtained, which could be used in POS-1 translational regulation studies *in vivo* or *in vitro*. However, there are thousands of 3' UTRs that are possibly regulated by POS-1, along with thousands of other combinations of RNA-binding protein – maternal mRNA interactions. Library MosSCI could reduce the time it takes to perform these studies by generating the thousands of necessary transgenic strains required for these experiments. Library MosSCI also serves as a tool to create libraries of the same transgene with varying mutation sites. In this sense, one 3' UTR could be mutated in multiple binding sites for various RNA-binding proteins. Each mutated binding site would correlate to the desired study, whether it be, for example, POS-1 binding to a single site, POS-1 binding to two closely related sites, POS-1 binding alongside GLD-1, POS-1 binding at one/two sites as opposed to other close proximity sites. Instead of injecting a library of different 3' UTRs, one could inject a library of the same 3' UTR with any range of mutations. With

further optimization, Library MosSCI can become the leading method of transgenic *C. elegans* strain generation for use in understanding the mechanics of translational regulation, especially in germline progenitor cells.

BIBLIOGRAPHY

Bessereau JL, Wright A, Williams DC, Schuske K, Davis MW, Jorgensen EM (2001) Mobilization of a *Drosophila* transposon in the *Caenorhabditis elegans* germ line. *Nature* 413(6851): 70-74.

Carmel AB, Wu J, Lehmann-Blount KA, Williamson JR (2010) High-affinity consensus binding of target RNAs by the STAR/GSG proteins GLD-1, STAR-2 and Quaking. *BMC Molecular Biology* 11: 48.

Farley BM, Pagano JM, Ryder SP (2008) RNA target specificity of the embryonic cell fate determinant POS-1. *RNA* 14: 2685-2697s.

Farley BM, Ryder SP (2008) Regulation of Maternal mRNAs in Early Development. *Critical Reviews in Biochemistry and Molecular Biology* 43: 135-162.

Frøkjaer-Jensen C, Davis MW, Hopkins CE, Newman BJ, Thummel JM, Olesen S, Grunnet M, Jorgensen EM (2008) Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nature Genetics* 40: 1375-1383.

Gu W, Shirayama M, Conte D Jr., Vasale J, Batista PJ, Claycomb JM, Moresco JJ, Youngman E, Keys J, Stotltz MJ, Chen CG, Chaves DA, Duan SE, Kasschau KD, Falgren N, Yates JR III, Mitani S, Carrington JC, Mello CC (2009) Distinct Argonaute-mediated 22G-RNA pathways direct genome surveillance in the *C. elegans* germline. *Molecular Cell* 32(2): 231-244.

Matranga C, Ryle AM (2010) Double-stranded RNA-dependent ATPase DRH-3. The *Journal Biological Chemistry* 285(33): 25363-25371.

Merritt C, Rasoloson D, Ko D, Seydoux G (2008) 3' UTRs are the primary regulators of gene expression in the *C. elegans* gremline. *Current Biology* 18: 1476-1482.

Ogura K, Kishimoto N, Mitani S, Gengyo-Ando K, Kohara Y (2003) Translational control of maternal *glp-1* mRNA by POS-1 and its interacting protein SPN-4 in *Caenorhabditis elegans*. *Development* 130: 2495-2503.

Pagano JM, Farley BM, Essien KI, Ryder SP (2009) RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. *Proceedings of the National Academy of Sciences of the United States of America* 106(48): 20252-20257.

Riddle DL, Blumenthal T, Meyer BJ, Priess JR. C. Elegans II. 1997. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Rieckher M, Kourtis N, Pasparaki A, Tavernarakis N (2009) Transgenesis in *Caenorhabditis elegans*. *Methods in Molecular Biology* 561: 21-39.

Sulston JE, Schierenberg E, White JG, Thomson JN (1983) The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Developmental Biology* 100: 64-119.

Tabara H, Hill RJ, Mello CC, Priess JR, Kohara Y (1999) *pos-1* encodes a cytoplasmic zinc-finger protein essential for germline specification in *C. elegans*. *Development* 126: 1-11.

Zhang C, Montgomery TA, Gabel HW, Fischer SE, Phillips CM, Fahlgren N, Sullivan CM, Carrington JC, Ruvkun G (2011) *mut-16* and other *mutator* class genes modulate 22G and 26G siRNA pathways in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America* 108(4): 1201-1208.