Upstream Elements Regulating *mir-241* and *mir-48*

Hanna Vollbrecht, Tamar Resnick, and Ann Rougvie
University of Minnesota: Twin Cities
Undergraduate Research Scholarship 2012-2013

Abstract

Caenorhabditis elegans is a free-living nematode that has been extensively used to study developmental timing, an evolutionarily conserved and important process that when disrupted can contribute to the etiology of some congenital human diseases. While proteins are known to be important to developmental timing, studies in recent years have also indicated the importance of short, non-coding RNAs, called microRNAs, which control gene expression. *mir-241*, *mir-48,* and *mir-84* are microRNAs that all belong to the *let-7* family of microRNAs, which act redundantly to influence development in *C. elegans*. A mutant phenotype of these microRNAs suggests that they function in the lateral hypodermal seam cells. *mir-241* and *mir-48* reside together on chromosome V and seem to be under the control of some of the same regulatory sequences located upstream of *mir-241*, the "*mir-241* promoter". The goal of this project was to identify and characterize a specific seam cell enhancer element within this promoter. We hypothesized that a subfragment of a 750 bp sequence drives *mir-241/mir-48* expression in the seam cells. The most consistent seam cell expression was seen from a 275 bp fragment, thus suggesting this promoter element is important in directing *mir-241/mir-48* seam expression.

Introduction

Organisms have evolved molecular mechanisms that ensure developmental events occur at the proper time and in the appropriate sequence. Developmental timing mechanisms coordinate the growth of an organism and influence many processes that are essential to the diversity of life and the proper functioning of the organism, and thus are an important area of research. Changes in developmental timing contribute to significant evolutionary change, and a large amount of evolutionary variation is attributable to simple changes in genes important to developmental timing, called heterochronic genes (Gould, 1992). Additionally, defects in developmental timing can result in damaging birth defects (Wilson, 1988). Previous studies have provided evidence that mechanisms of developmental timing are highly conserved, thus research in model organisms concerning development can provide valuable insight into the development of humans (reviewed in Resnick *et al.*, 2010)

The nematode *Caenorhabditis elegans* is a useful organism for studying the timing of developmental events, due to its simple anatomy, fast growth rate and essentially invariant cell lineage (Sulston and Horvitz, 1977). There are five postembryonic developmental stages that are labeled L1, L2, L3, L4, and adult. These stages are separated by four sequential molts. Development in *C. elegans* is defined by a set cell lineage, characterized by some stage-specific developmental events. Mutations in heterochronic genes cause temporal changes in these developmental events, resulting in their execution at a different time than usual. Two types of timing mutants are defined, precocious and retarded. Precocious mutants skip an event in development, and thus subsequent events happen too early. Retarded mutants repeat a developmental event, and thus subsequent events occur too late in the developmental timeline (Ambros and Horvitz, 1984). Through identification of timing defects caused by mutations in certain genes, heterochronic phenotypes can be characterized and the affected genes subsequently cloned and analyzed, revealing mechanisms of developmental timing in *C. elegans*.

While many proteins are known to be of great importance to developmental timing and some have been previously identified in *C. elegans*, studies in recent years have also demonstrated the importance of microRNAs (miRNAs) in such events. miRNAs are short non-coding RNA molecules, approximately 22 bp in size. These molecules typically bind to target sites of antisense complementarity in the 3' UTRs of mRNA molecules, generally resulting in lower levels of translation or degradation of mRNA (Lee *et al.*, 1993; Reinhart *et al.*, 2000). miRNAs are thus negative regulators of gene expression.

mir-48, *mir-241,* and *mir-84* are members of the *let-7* family of miRNAs and act redundantly to control developmental time. Previous research suggests that these miRNAs function together to control the L2-to-L3 larval transition. A knockout of *mir-48* alone results in little to no developmental defects. A double knockout of *mir-48*/*mir-241,* however, results in defects in the timing of seam cell development (Abbott *et al.*, 2005). Seam cells are hypodermal cells that are found along the left and right side of the worm's body. They are useful for studies of temporal development because they undergo two distinct events at specific times. One is seam cell proliferation, an event that occurs in the second larval stage, in which the number of seam cells increases from 10 to 16. The other is the formation of ridges called alae on the cuticle made by the seam cells, which occurs during the molt from the fourth larval stage to the adult stage (reviewed in Resnick *et al.*, 2010). In animals with a double knockout of *mir-48* and *mir-241,* an abnormally large number of seam cells are present in the L3 stage, suggesting a retarded phenotype that can be attributed to repetition of the L2-stage seam cell proliferation program during the L3 stage. Further, double mutant animals display incomplete adult-stage alae indicating some cells have retained their larval characteristics. All of these phenotypes are enhanced in triple mutant *mir-48/ -84/ -241* worms (Abbott *et al.,* 2005). These mutant phenotypes indicate the importance of these miRNAs in temporal development, particularly of *mir-48* and *mir-241*, and suggest that they likely have redundant functions. In this project, we asked specific questions about the expression profiles of these miRNAs and aimed to analyze upstream regulatory sequences of *mir-48* and *mir-241*, focusing on a seam enhancer.

Materials and Methods

Traditional cloning methods were used to design plasmids for injection into *C. elegans* to create transgenic animals. Plasmids pPD97.78 and pPD95.67 were cut with appropriate restriction enzymes (NEB), and the sequences of interest were amplified from genomic wild-type (N2) *C. elegans* DNA by PCR (iProof Hi-Fidelity DNA Polymerase), restriction enzyme digested (NEB), and ligated into the vector using T4 DNA Ligase. Plasmids were analyzed by restriction digest (enzymes from NEB, Promega) to establish that their restriction fragment lengths were consistent with intended plasmid design. Six plasmids (Table 1) were constructed.

Table 1. Plasmid Names with corresponding insert and the plasmid backbone utilized. Worm lines were named according to the plasmid with which they were injected. See figure 1.

Plasmid Name	Plasmid Backbone	Insert	Worm Lines
pHV1	pPD95.67	Fragment 1	HV1-3K, HV1-Is4a4, HV1-Is4j4, HV1-2.12-4, HV1-2.12-5
pHV2	pPD95.67	Fragment 2	HV2-A1
pHV3	pPD95.67	Fragment 3	HV3-9-3
pHV4	pPD97.78	Fragment 1	HV4-I-9, HV4-H10
pHV5	pPD97.78	Fragment 2	HV5-9.6, HV5-1.12, HV5-3.3
pHV6	pPD97.78	Fragment 3	HV6-6-2, HV6-9-1, HV6-9-2

Once the six plasmids were characterized, injection mixes were made, each with the appropriate transgene plasmid and *Pmyo-2::tdtomato*, a plasmid that results in fluorescent red expression in the pharynx and is used as a marker for transgenic animals. These mixes were injected into N2 wild type worms at a concentration of 30 ng/ul for the transgene and 1 ng/ul for *Pmyo-2::tdtomato*. Worms injected with plasmids spontaneously arrange the plasmid DNA into extrachromosomal arrays that contain many copies of each plasmid injected. Because worms are holocentric, they maintain these extrachromosomal arrays rather well, though they are lost at some frequency, leading to mosaicism in the transgenic animals. The goal was to create three independent worm lines for each plasmid, a total of 18 lines, to ensure that patterns are similar from independent arrays.

Two integrated lines were made, called HV1-Is4a4 and HVI-Is4j4. These worms were created by UV-irradiating HV1-3K worms containing extrachromosomal arrays. The UV light causes double-stranded breaks in the chromosomes, and these are sometimes repaired to incorporate the extrachromosomal arrays into the chromosome. This ensures that all cells and all worms contain the arrays; however, overexpression does occur and silencing can happen, thus multiple lines must again be considered.

Results

mir-241 and *mir-48* are located about 1.8 kilobases (kb) apart in the genome and seem to be under the control of some of the same regulatory sequences located upstream of *mir-241*, which I will hereafter refer to as the "*mir-241* promoter" (Figure 1a). Previous work in the Rougvie lab has determined that 4.5 kb of *mir-241* promoter results in all normal expression and function of *mir-241/-48*. On the other hand, 2 kb of *mir-241* promoter facilitates many aspects of *mir-241/-48* expression, with the exception of proper seam cell development. This suggests a *cis*-regulatory element required for proper seam cell development is located upstream of the 2kb promoter. Moreover, a previous researcher determined that a 750 basepair (bp) sequence just upstream of the 2 kb non-rescuing promoter (Figure 1a, star) is sufficient to direct expression of a *gfp* reporter in seam cells (D. Hernanadez, personal communication). In this study, we hypothesized that the 750 bp sequence contains an enhancer that drives expression of the miRNAs in the seam cells and is meaningful for their function. Characterizing this element will allow us to better understand the regulation of miRNAs in the seam. Additionally, this work could be useful in identifying *trans*-acting regulators of these miRNAs, which would be valuable in decoding heterochronic gene pathways.

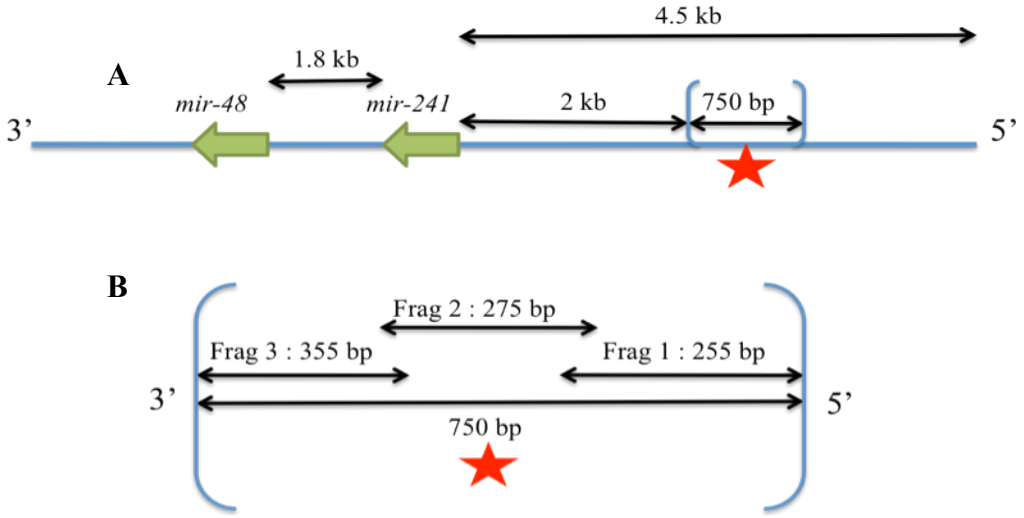

Figure 1. (a) Genomic representation of the *mir-241* and *mir-48* genes in *C. elegans*, with surrounding sequences included. Not to scale. (b) Fragmentation of 750 bp sequence of interest starred in (a). Not to scale

To test whether a seam cell specific enhancer could be identified, the 750 bp sequence was split into three overlapping pieces, which are approximately 250-300 bp, but not identical in size, named as fragment 1, 2, and 3 (Figure 1b). Using the BLAST program (Altschul *et al.*, 1990) conserved sequences that are found in other parts of the genome of *C. elegans* and in related nematodes were identified within this 750 bp sequence. In selecting the three pieces to split this sequence into, conserved elements were maintained together. Each of the three subfragments was inserted into a reporter plasmid containing the green fluorescent protein (*gfp*) gene. As the DNA pieces being inserted were of a small size, it was possible that they would not serve as fully functional promoters on their own. Thus, for each subfragment, two plasmids were used, pPD97.78 and pPD95.67. pPD95.67 contains a *gfp* gene with no promoter upstream of it, while pPD97.78 contains the same *gfp* gene with a *pes-10* minimal promoter upstream. The *pes-10* minimal promoter produces little to no *gfp* expression on its own, and is commonly used to test fragments for the presence of enhancers.

These plasmids were injected into wild type *C. elegans* with *Pmyo-2::td-tomato*, a plasmid that results in expression of a red fluorescent protein in the pharynx when transgenic animals are successfully made. Transgenic animals with pharyngeal red fluorescent protein (*rfp*) expression were then picked and imaged using high power fluorescence microscopy to identify *gfp* expression patterns. GFP was analyzed to gain insight into the expression pattern of the *mir-241* promoter, both spatially and also over developmental time.

In imaging, we selected transgenic worms at each developmental stage as determined by their gonad morphology and identified the tissues in which *gfp* expression was seen. Mosaicism was a common theme in all lines, as different worms of the same stage expressed *gfp* to a different degree in certain tissues, and some would not express *gfp* at all. A lack of a signal does not necessarily mean that the transgene would not be expressed in a certain tissue, just that the transgene may have been lost from those cells. Thus a lack of expression does not necessarily conclude a lack of expression as promoted by the specific promoter element until a statistically significant number of worms have been examined. Results are summarized in Table 2 and representative images of characteristic expression patterns are shown in figures 2-5. Images are representative of common expression patterns, though they do not capture all the variability within the strains.

Table 2. Expression patterns and morphological characteristics of *mir-241* promoter elements driving *gfp* reporter gene.

Transgene	Tissue Specific GFP Expression	Vulval Blip?	Total Worms Imaged
pHV1	No tissue specific expression. Could be due to array silencing. More imaging needed.	No	33
pHV2	Gonad somatic sheath (L2, L3, L4) Seam cell expression (L2, L3)	Yes	30
pHV3	Not characterized, likely similar to pHV6	No	6
pHV4	Neuronal expression (L2)	No	10
pHV5	Gonad somatic sheath (L2, L3, L4) Seam cell expression (L2, L3) Tail expression (L1, L2, L3, L4)	Yes	17
pHV6	Tail expression (L1, L2, L3, L4)	No	12

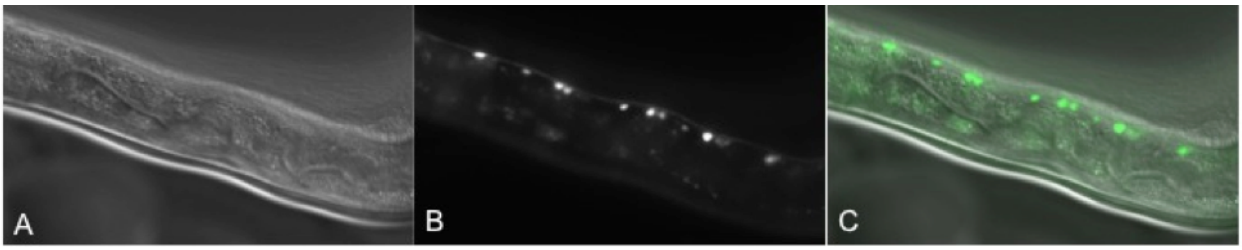

Figure 2. (A) DIC image of L2 HV4 transgenic worm (B) GFP expression is visualized along the ventral side of the animal in the neurons (C) GFP and DIC merge

HV1 transgenic worms lacked any significant tissue-specific GFP expression. HV4 worms, which contain the same promoter element as HV1 worms – fragment 1, also generally did not show any significant tissue-specific GFP expression, although some neuronal expression was observed in some L2 worms (Fig 2). This may suggest that fragment 1 lacks regulatory elements for gene expression or could be due to array silencing in these lines. Thus, creating more transgenic worm lines along with more imaging is necessary to effectively characterize the expression profile of fragment 1.

HV2 transgenic worms most often showed GFP expression in the somatic gonad, including sheath cells, across L2, L3, and L4 stages (Fig 3G-3I). However, some HV2 worms showed strong seam cell expression, particularly during the L2 and L3 stages (Figure 3A-F). HV5 worms, which contain the same promoter element as HV2, generally demonstrated expression similar to HV2 worms, as GFP expression was noted in the seam cells and in the somatic gonad (Figure 4A-C). The most common expression was again in the gonad cells, particularly in the somatic sheath, across L2, L3, and L4 stages (Figure 4D-F). Seam cell expression was noted again during the L2 and L3 stages. A new expression pattern was seen in the tail region of the worm in the L1, L2, L3, and L4 stages (Figure 4G-4I). The seam expression seen from the pHV2 and pHV5 transgene suggests that Fragment 2 contains a seam cell enhancer.

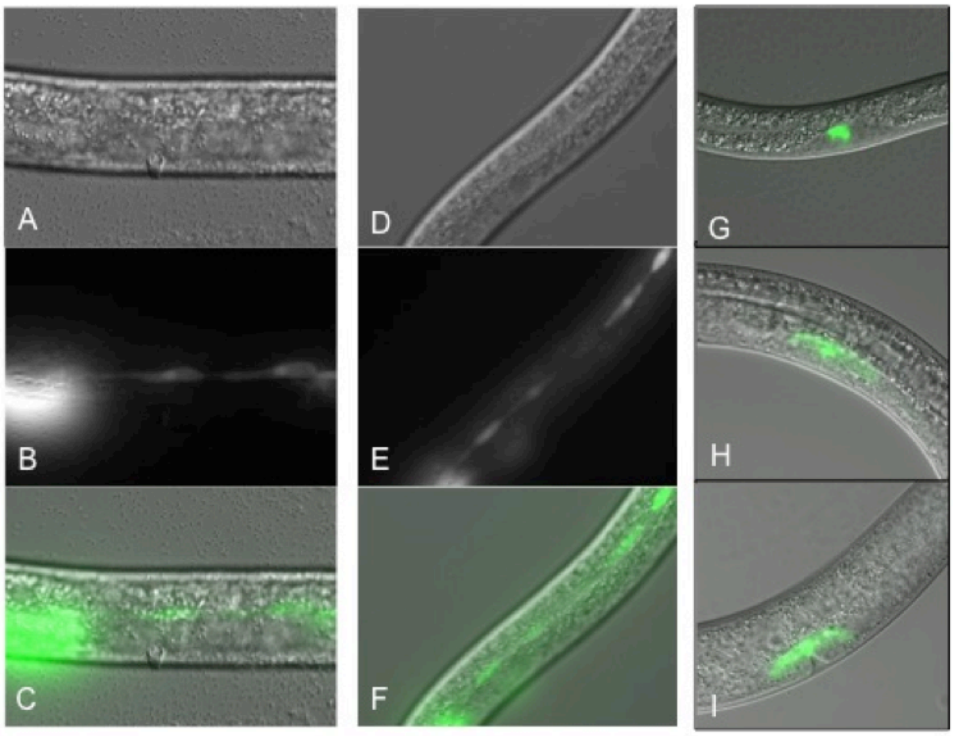

Figure 3. pHV2 confers GFP expression in seam cells. (A,D) DIC; (B,E) GFP fluorescence; (C,F,G,H,I) Merge. (A-C) L2 transgenic worm displaying GFP expression in the seam cells (D-F) L3 transgenic worm displaying GFP seam expression. (G-I) Somatic gonad GFP expression in HV2 transgenic worms at different larval stages. (G) L2 transgenic worm (H) L3 transgenic worm (I) L4 transgenic worm.

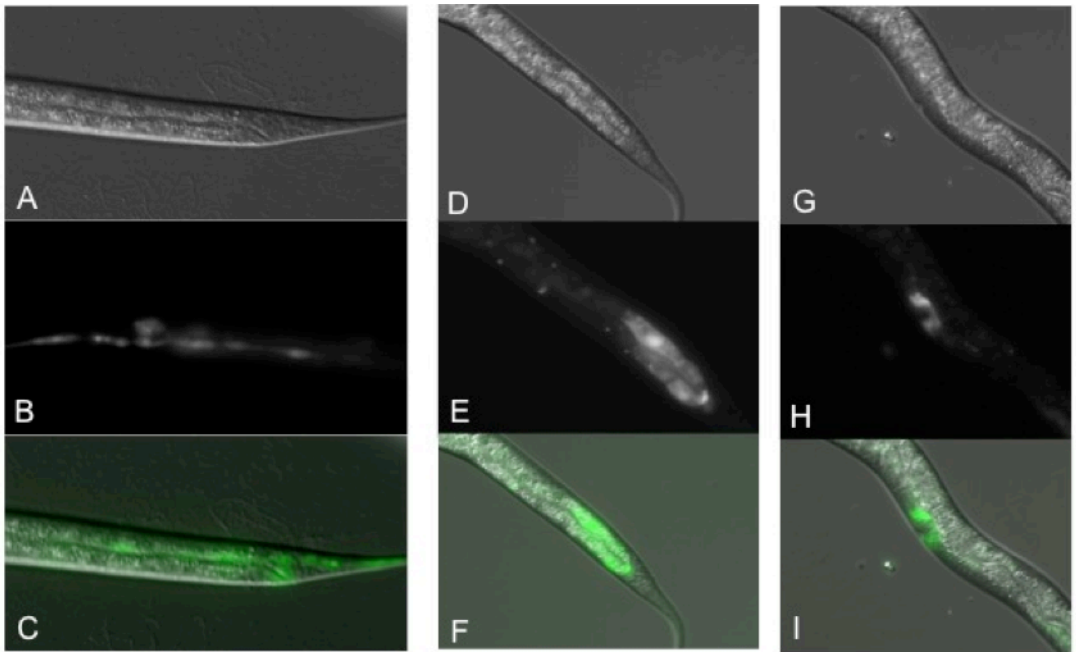

Figure 4. pHV5 produces GFP expression in several different tissues. (A,D,G) DIC; (B,E,H) GFP fluorescence; (C,F,I) Merge. (A-C) L1 transgenic worm displaying GFP expression in seam cells. (D-F) L2 transgenic worm with fluorescence in tail-end of animal (G-I) L4 transgenic animal with GFP expression in gonadal somatic sheath cells

Few HV3 worms were characterized and it was difficult to draw conclusions off the preliminary data gathered; however, it is likely that the expression is similar to that seen in the HV6 transgenic worms. HV6 transgenic worms displayed the same tissue specific GFP expression in the tail region of the worm as seen in pHV5 worms, in the L2, L3, and L4 stages (Figure 5). This was the only significant pattern noted in HV6 transgenic worms. pHV5 and pHV6 contain different promoter pieces, thus this may indicate that there is a repetitive promoter element in both Fragment 2 and Fragment 3 that directs this tail expression.

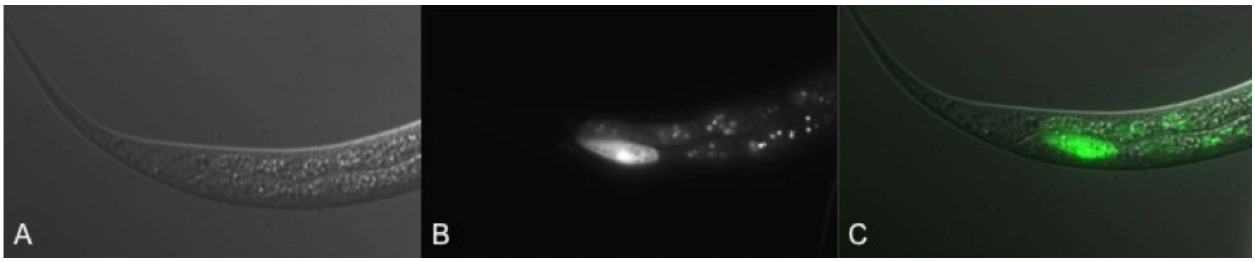

Figure 5. (A) DIC image of HV6 L2 transgenic worm. (B) GFP expression seen in tissue at tail end of animal. (C) Overlay of GFP expression on DIC image. All HV6 transgenic worms viewed that had tissue-specific GFP expression looked similar to this worm.

A vulval blip phenotype was also seen solely in the pHV2 and pHV5 worm lines. However, not all transgenic adult worms of these lines exhibited this phenotype. A developmental disruption is an unexpected result as the transgenes are only expressing GFP. Likely, the high copy number of pHV2 and pHV5 elements in these transgenic worm lines is competing with another gene that influences vulval development for a *trans*-acting factor. Thus,

the expression of this other gene as promoted by a specific transcription factor may be decreased, resulting in vulval developmental disruption.

Discussion

mir-48/mir-241 encode microRNAs that are important to time stage-specific seam cell lineage patterns. The work reported here has narrowed down the location of a seam cell enhancer upstream of these miRNAs to a 275 bp sequence, labeled as fragment 2. This piece of promoter in transgenes pHV2 and pHV5 resulted in the most consistent GFP seam cell expression in transgenic worms, leading to the suggestion that this segment may be sufficient for promoting expression of *mir-48* and *mir-241* in the seam. The variable expression seen, specifically expression in some but not all seam cells and no expression in seam cells in some transgenic animals, could be due to mosaicism from the extrachromosomal arrays. By integrating these extrachromosomal arrays, we can eliminate some of this mosaicism for a clearer picture of fragment 2 expression.

Variability could also be attributed to the fact that we may have disrupted an essential promoter element in splitting up the 750 bp "promoter" region, and thus could test different sized pieces with more overlap to effectively characterize the enhancer elements. Another approach to this would be to look side by side at worms containing HV2 (or HV5) and worms containing the 750 bp piece we started with. If the seam expression looked similar then we likely narrowed down the 750 bp fragment to the correct segment.

We originally expected to see more seam cell expression in the pHV5 worm lines, as expression of GFP should have been enhanced in comparison to pHV2 worm lines due to the presence of the *pes-10* minimal promoter. This suggests that a transcription site exists in the 275 bp piece (fragment 2) or elsewhere on the array that can be used for transcription of the GFP, thus making the *pes-10* promoter unnecessary. This lack of enhancement is curious, but may be attributable to mosaicism in the population of worms imaged, a question that could be answered by integrating the array to see if variability decreases.

Further research may be directed toward additional imaging of the transgenic worm lines to more fully characterize their GFP expression and examine the dynamics of expression over time. If more worms are characterized across all larval stages, a more precise understanding of the patterns of expression *mir-48* and *mir-241* over developmental time may be possible. Another useful experiment would be to create integrated lines with all plasmids, particularly with pHV2 and pHV5 plasmids as these are the plasmids of the most interest due to the seam expression seen. This would help in minimizing mosaicism.

The 750 bp sequence could also be split into different pieces to see if this changes expression patterns, or the three sequences of interest examined in this experiment could be split into even smaller pieces. If we carried out this smaller portioning, the likely direction would be to split the second fragment (HV2, HV5) into smaller pieces as it resulted in the most consistent GFP expression. In this way, we could more specifically characterize the *cis* regulatory element(s) that are directing seam expression of *mir-48* and *mir-241*. These *cis* regulatory elements that direct seam expression may be similar to other *cis* elements that direct seam expression of other genes in the *C. elegans* genome. Knowing the *cis* elements as precisely as possible will help in knowing what to look for by other genes. Additionally, a lot of different tissues were green in HV6 lines, and thus smaller fractioning of fragment three could be valuable in separating out some of these patterns as being directed by different *cis* elements.

Additionally, the *trans* regulators of *mir-241* and *mir-48* could be characterized, which would be useful in uncovering developmental timing pathways. This could be done via RNAi experiments in which plasmids are ingested by worms and make dsRNAs that are then chopped into small RNAs in their own cells. These RNAs block the expression of certain transcription factors of interest. Experiments would likely be carried out with pHV2 or pHV5 transgenic worms to determine if expression changes when the action of a certain transcription factor is inhibited. It will be essential to establish lines with little to no mosaicism prior to performing any RNAi experiments. To be able to note a change in GFP expression post-RNAi that could be attributed to the inhibition of a *trans* regulator, the GFP expression must be fairly consistent in all transgenic worms.

Developmental timing in *C. elegans* is important to understand, due to the highly evolutionarily conserved nature of development across many species. This research allowed further characterization of the upstream promoter region of *mir-241* and *mir-48* and suggests that a 275 bp section of promoter is sufficient to direct expression. While further research is necessary to understand the mechanism of transcription of these genes, the results from this experiment lead us closer to the eventual goal of identifying the key elements in regulating *mir-48* and *mir-241*. Identification of such enhancers may reveal regulatory elements that are essential to expression of many different microRNAs, as well as further our understanding of the development of *C. elegans*.

References:

Abbott, AL, Alvarez-Saavedra, E, Miska, EA, Lau, NC, Bartel, DP, Horvitz, HR, Ambros, V. "The *let-7* microRNA family members *mir-48*, *mir-84*, and *mir-241* function together to regulate developmental timing in *Caenorhabditis elegans*." Dev. Cell 2005, 9: 403-414.

Ambros, V, Horvitz, RH. "Heterochronic Mutants of the Nematode *Caenorhabditis elegans*." Science 1984, 226: 409-416.

Altschul, SF, Gish, W, Miller, W, Myers, EW, Lipman, DJ. "Basic local alignment search tool." J. Mol. Biol. 1990, 215: 403-410.

Gould, SJ. "Ontogeny and phylogeny—revisited and reunited." Bioessays 1992, 14: 275-279.

Lee, RC, Feinbaum, RL, Ambros, V. "The *C. elegans* Heterochronic Gene *lin-4* encodes small RNAs with antisense complementary to *lin-14*." Cell 1993, 75: 843-854.

Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G. "The 21- nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans." Nature 2000, 403: 901–906.

Resnick, TD, McCulloch, KA, Rougvie, AE. "miRNAs Give Worms the Time of Their Lives: Small RNAs and Temporal Control in *Caenorhabditis elegans*." Developmental Dynamics 2010, 239: 1477-1489.

Rougvie, AE. "Control of Developmental Timing in Animals." Nature Review Genetics 2001, 2(9): 690-701.

Sulston, J.E. & Horvitz, H.R. (1977) "Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*" *Developmental Biology* 56: 110-156.

Wilson, G.N. (1988) "Heterochrony and human malformation" *American Journal of Medical Genetics* 29:311-321.