UPPSALA
UNIVERSITET

Investigating the expression and function of RNA-binding protein FUBL-1 in *C. elegans*

Védís Mist Eyju Agnadóttir

Degree project in biology, Master of science (2 years), 2023
Examensarbete i biologi 60 hp till masterexamen, 2023
Biology Education Centre and Department of Cell and Molecular Biology, Uppsala University
Supervisor: Andrea Hinas

Abstract

RNA interference (RNAi) is an important mechanism of gene silencing in *Caenorhabditis elegans*, in which short RNAs direct sequence-specific silencing of gene expression mediated by Argonaute proteins. RNAi in *C. elegans* can be sorted into exogenous RNAi, where the short RNAs originate from foreign RNA sequences, and endogenous RNAi, where the short RNAs originate from RNA sequences in the genome. One endogenous RNAi pathway is the ERGO-1 pathway, which is active in the germline and in embryos. Mutants deficient in the ERGO-1 pathway show an increased response to exogenous RNAi, which is thought to be due to competition between exogenous and endogenous silencing pathways.

FUBL-1 is an RNA-binding protein found in *C. elegans*, which has three predicted functional isoforms: isoforms a, b and c. Prior research by the Hinas group has indicated that FUBL-1 may play a role in the ERGO-1 pathway of RNAi. FUBL-1 deletion mutants show an increased response to exogenous RNAi similar to ERGO-1 mutants, and also an upregulation of ERGO-1 target genes. They have also found that FUBL-1 is broadly expressed in somatic tissues, but germline expression has not been confirmed. The function of the three different isoforms has not yet been examined.

The aim of this study was to further investigate the function of FUBL-1 by assessing the function of the different isoforms through RT-qPCR of *C. elegans* with mutations affecting the different isoforms, to use immunofluorescence staining to see whether FUBL-1 is expressed in the germline, and to identify FUBL-1 RNA targets through CLIP-seq. Preliminary RT-qPCR results indicated that the upregulation of ERGO-1 targets is less in isoform mutants than in FUBL-1 deletion mutants, indicating partial redundancy of the isoforms. Immunofluorescence staining showed that FUBL-1 is expressed in the germline nuclei. Growth protocols have been optimized and crosslinking has been performed on worms, but the full CLIP-seq protocol has not been performed yet.

Table of Contents

Abbreviations

RNAi	RNA interference
dsRNA	Double-stranded RNA
sRNA	Small RNA
endo-RNAi	Endogenous RNAi
exo-RNAi	Exogenous RNAi
siRNA	Short interfering RNA
RdRp	RNA-dependent RNA polymerase
Eri	Enhanced RNAi
FUBL-1	Far upstream binding-like protein 1
FUBP	Far upstream binding protein
iso	Isoform
NGM	Nematode growth medium
PCR	Polymerase chain reaction
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
CLIP-seq	Crosslinking and immunoprecipitation sequencing
Ct	Cycle threshold
cDNA	Complementary DNA

Introduction

RNAi in C. elegans

RNA interference (RNAi) is a conserved biological process in most eukaryotes in which double-stranded RNA (dsRNA) molecules are processed to single-stranded small RNAs (sRNAs) (20-26 nucleotides) that direct the sequence-specific silencing of gene expression. The dsRNA molecules can originate from the organism's own genome (endogenous RNAi, endo-RNAi) or from external sources such as viruses (exogenous RNAi, exo-RNAi) (Hannon 2002). RNAi was first described in the nematode *Caenorhabditis elegans* (Fire *et al.* 1998), where it plays a vital role in gene regulation, including development and protection against threats to genome stability such as transposable elements and viruses (Almeida *et al.* 2019a).

RNAi in *C. elegans* is involved in mechanisms of transcriptional gene silencing as well as post-transcriptional gene silencing. Transcriptional gene silencing by RNAi is called nuclear RNAi as the silencing occurs in the nucleus through mechanisms such as histone modifications, such as H3K9 methylation which defines transcriptionally silent heterochromatin, or impeding the elongation of RNA transcripts by RNA polymerase. Post-transcriptional gene silencing by RNAi involves the degradation of mRNA transcripts in the cytoplasm, and is therefore called cytoplasmic RNAi (Shiu & Hunter 2017). Both nuclear and cytoplasmic RNAi pathways in *C. elegans* involve a primary small RNA signal which triggers the production of secondary 22 nucleotide-long RNA molecules with 5' G bias, termed 22G RNAs. The primary sRNA pathways can be divided into exogenous and endogenous pathways. In the exo-RNAi pathway, dsRNA from the environment is cleaved into short interfering RNAs (siRNAs) by the Dicer enzyme, and the siRNAs then trigger the synthesis and amplification of 22G RNAs by RNA-dependent RNA polymerase (RdRP). The endogenous pathways involve the transcription of endogenous primary sRNA molecules, which then trigger the production of 22G RNAs. The endogenous pathways can be divided into the 21U pathway where the primary sRNA molecule is 21 nt long with a 5' U bias, and the 26G pathway where the primary sRNA molecule is 26 nt long with a 5' G bias (Almeida *et al.* 2019a). In cytoplasmic RNAi, the secondary 22G RNAs bind Argonaute proteins which degrade mRNA targeted by the 22G RNA. In nuclear RNAi, the 22G RNAs bind the Argonautes NRDE-3 (in the soma) or HRDE-1 (in the germline), which transport the 22G RNA into the nucleus (Guang *et al.* 2008, Buckley *et al.* 2012). In the nucleus, the Argonautes interact with other nuclear RNAi components; the complex then binds nascent RNA transcripts that are complementary to the 22G RNA and impede their elongation by

RNA polymerase, and directs the deposition of repressive H3K9 methylation marks at targeted sites (Guang *et al.* 2010, Burton *et al.* 2011, Buckley *et al.* 2012).

The 26G RNAs are low in abundance and produced by the RdRP RRF-3 with assistance from the ERI complex in the germline. They form two distinct subpopulations based on which Argonaute protein they associate with; the ALG-3/4 sRNAs which bind ALG-3 and ALG-4 in the spermatogenic germline of L4 hermaphrodites and males, and the ERGO-1 branch sRNAs which associate with ERGO-1 in the oogenic germline and in embryos (Almeida *et al.* 2019a). The elimination of oogenic/embryonic 26G RNAs leads to an increased RNAi response to exogenous dsRNAs, termed an Enhanced RNAi (Eri) phenotype (Lee *et al.* 2006). The Eri phenotype is thought to be due to competition between exogenous and endogenous RNAi pathways as exogenous and endogenous RNAi pathways have factors in common.

RNAi inheritance in C. elegans

RNAi in *C. elegans* is inherited from parent to progeny, with parental RNAi induction leading to continued silencing in future generations (Fire *et al.* 1998). The duration of this effect can range from a few generations to dozens of generations (Vastenhouw *et al.* 2006, Rankin 2015). Epigenetic changes occur when an organism is exposed to a stimulus which leads to alterations in gene expression in some (or all) of its cells, and these alterations are passed onto the daughter cells of said cells, even in the absence of the trigger stimulus, despite no genetic change occurring. Epigenetic inheritance occurs when the descendants of an organism exposed to a stimulus also show the epigenetic changes (Lacal & Ventura 2018). This inheritance may be a way for organisms to pass information about their environment to their progeny, thus giving them a selective advantage. As it is not accompanied by any genetic changes, RNAi inheritance is a form of epigenetic inheritance.

As described above, in Eri mutants, the elimination of oogenic/embryonic 26G RNAs leads to an increased exo-RNAi response. Worms with a heterozygous *eri*/+ genotype do not show an Eri phenotype, and 25% of their offspring have a homozygous *eri/eri* genotype. The Eri phenotype shows a strong maternal rescue, meaning the Eri phenotype penetrance in the offspring of *eri*/+ heterozygotes is lower than the expected 25% (Zhuang & Hunter 2011, Almeida *et al.* 2019a). This indicates that maternal contribution plays an important role in the 26G silencing pathway function. The maternal effect appears to be specific to mutants where the ERGO-1 pathway 26G RNAs are affected. Zygotic 26G RNAs can induce the production

of 22G RNAs in the absence of maternal 26G RNAs, indicating that while maternal 26G RNAs are sufficient for proper RNAi response, they are not necessary (Almeida *et al.* 2019a).

FUBL-1

Far upstream element binding protein like 1 (FUBL-1) in *C. elegans* is an ortholog of human far upstream element binding proteins 1 and 3 (FUBP-1 and FUBP-3) and is predicted to possess mRNA binding activity and be involved in mRNA processing and alternative splicing. The *fubl-1* gene is predicted to have three coding sequences: isoform a, isoform b1/b2, and isoform c (WormBase web site, https://wormbase.org/species/c_elegans/gene/WBGene00007534, release WS288, 2023-05-10) (Figure 1). Isoform a, the most common isoform, contains a predicted nuclear localization signal (NLS), while isoforms b and c do not. A fourth isoform is also found, isoform d, which has an open reading frame of only 20 amino acids and has been annotated as a non-coding RNA. FUBL-1 has been found to localize both in the nucleus and the cytoplasm, and it is possible that switching between isoforms controls the localization of FUBL-1, though it is also possible that the localization is controlled through other means such as phosphorylation (Roy *et al*., manuscript in preparation). FUBL-1 has been shown to interact with proteins involved in pre-mRNA processing, post-transcriptional modification, and chromatin remodeling (Li *et al.* 2004, Wu *et al.* 2017). FUBL-1 has also been associated with small RNA pathways through patterns of phylogenetic conservation (Tabach *et al.* 2013), and knockdown of FUBL-1 has been shown to lead to disruption of exo-RNAi (Kim *et al.* 2005).

Figure 1. A schematic of the predicted fubl-1 *isoforms b1, b2, a, c and d. Purple indicates the 5' and 3' untranslated regions, turquoise indicates exons and thin black lines indicate introns. Isoform d is gray as it is predicted to be non-functional* (WormBase web site, https://wormbase.org/species/c_elegans/gene/WBGene00007534, release WS288, 2023-05-10).

Prior research by the Hinas group indicates a role of FUBL-1 in the 26G ERGO-1 RNAi pathway, as well as inheritance of RNAi. While RNAi knockdown of FUBL-1 led to decreased GFP silencing efficiency in a GFP knockdown assay, worms with a *fubl-1* loss-of-

function (*fubl-1Δ*) mutation show an increased sensitivity to exo-RNAi in the same assay. The increased exo-RNAi in *fubl-1Δ* worms appears to be primarily due to an increase of inherited RNAi, rather than same-generation RNAi, which indicates that FUBL-1 may play a role in RNAi inheritance. ERGO-1 target genes are upregulated in *fubl-1Δ* mutants, indicating a role of FUBL-1 in ERGO-1 26G silencing (Roy *et al*., manuscript in preparation).

So far research has indicated that FUBL-1 plays a role in ERGO-1 pathway 26G silencing, however more detail about the contribution of FUBL-1 to RNA silencing and the RNAs it interacts with remains unknown. One aim of this study was to assess the roles of the different FUBL-1 isoforms in the function of FUBL-1 through RT-qPCR of ERGO-1 target genes in different FUBL-1 mutants. Due to problems with reverse transcription the qPCR was only performed on one biological replicate, but results from said biological replicate indicate that the different FUBL-1 isoforms may be partially redundant and furthermore, isoform d may be functional despite what has previously been thought. Prior results indicate that FUBL-1 is related to RNAi inheritance, and immunofluorescence staining in this study showed that FUBL-1 is expressed in the germline. Crosslinking was performed on 100,000 *C. elegans* early adults and protein was extracted from them with the intent of performing CLIP-seq, however immunoprecipitation has not been performed yet due to experimental difficulties and time constraints.

Materials and Methods

Worm strains and maintenance

Worm strains N2 Bristol (Brenner, 1974), AHS205 *fubl-1::3X FLAG* (6x outcrossed), AHS171 (*fubl-1*(gk668481) V (*fubl-1* all iso(-)), 3x outcrossed), AHS158 *fubl-1*(tm2769) V (4x outcrossed) (Roy et al., manuscript in preparation), VC40712 *fubl-1*(gk773341) V (not outcrossed), AHS170 *fubl-1*(gk660081) V (4x outcrossed) (Hinas, unpublished) were obtained from Hinas lab. Worm strains AHS207 (*fubl-1*(tm2769) V (*fubl-1* deletion), 6x outcrossed), AHS208 (*fubl-1*(gk660081) V (*fubl-1* iso-A(-)), 6x outcrossed), AHS209 (*fubl-1*(gk773341) V (*fubl-1* iso-B(-)), 6x outcrossed), AHS210 (*fubl-1*(gk668481) V (*fubl-1* all iso(-)), 6x outcrossed) were obtained during this project through outcrossing of strains listed above.

All worms were grown at 20°C on nematode growth medium (NGM) plates seeded with OP50-strain *E. coli* (Brenner 1974), except for high growth protocol worms for protein extraction. In the high growth protocol, worms were maintained on 8P-enriched peptone

plates seeded with NA22-strain *E. coli* (Evans 2006). Male N2 worms were induced by heat shocking L4 hermaphrodites at 30°C for 5 hours.

Genotyping

Single-worm lysis was performed on each worm by placing the worms in 2.5 µL 10 mM Tris-Cl (pH 8.5) with 120 µg/mL proteinase K. The reaction mixture with the worm was incubated for 10 minutes at -80°C before thawing at room temperature. Subsequently the reaction was incubated in a PCR machine for 1 hour at 65°C and 15 minutes at 95°C to inactivate the protease. The worm lysates were then used for genotyping PCR.

Genotyping PCR was performed using DreamTaq DNA polymerase (Thermo Scientific). Reaction volumes were 25 µL (2.5 µL worm lysate, 1X DreamTaq buffer, 0.2 mM dNTP, 0.1 µM forward primer, 0.1 µM reverse primer). Reaction conditions were 30 seconds denaturation at 95°C followed by 35 cycles of 30 seconds denaturation at 95°C, 30 seconds annealing at 51°C and 1 minute elongation at 72°C, followed by a final 5 minutes elongation period at 72°C. The primer pairs used for each isoform are listed in table S1.

Bleach synchronization

Worms were collected in 1X M9 buffer from mixed-stage plates of worms with a high number of gravid adults and embryos. The worms were then washed 3 times in M9 to remove bacteria before a bleach solution containing 1% sodium hypochlorite and 0.5M KOH was added to the worm pellet. The worms were then vortexed frequently and the bleach reaction went on for a maximum of 5 minutes until the worm bodies were disrupted and the reaction was stopped by adding 1X M9. The samples were then centrifuged and washed 3 times with 1X M9 to remove all residual bleach. A sucrose flotation was then performed on the pellets to remove any debris of dead worm bodies from the egg prep by suspending the pellet in 30% sucrose solution and centrifuging at 450 g for 6 minutes in a swing-bucket rotor. The upper 2 mL phase containing eggs was collected to a new tube, and sterile water was added to wash away the sucrose. The pellet was then resuspended in 1X M9 and the eggs were distributed on clean NGM plates. For RNA extraction, 500 eggs were distributed on each plate and for crosslinking and protein extraction 16,700 eggs were distributed on each plate. The following day, L1 larvae were collected in 1X M9 and distributed on either NGM plates with OP50 *E. coli* for RNA extraction or on 8P-enriched plates with NA22 *E. coli* for crosslinking and protein extraction.

RNA extraction

Synchronized adult worms were collected in 1X M9 and centrifuged to pellet. The pellet was washed twice with 1X M9 and the pellet was then frozen in liquid nitrogen for a minimum of 5 minutes. The samples were immediately ground with a micro pestle upon removal from liquid nitrogen before Trizol was added. After Trizol addition, samples were vortexed for 10-15 seconds and 0.2 volumes of chloroform were added. Samples were vortexed and then incubated for 10 minutes at room temperature before being centrifuged for 15 minutes at 16,200 g at 4°C. The aqueous phase was transferred to a new tube and a second chloroform extraction was performed in the same way. After extraction, an equal volume of isopropanol was added to the aqueous phase and the samples were incubated at room temperature for 10 minutes before centrifugation for 15 minutes at 16,200 g at 4°C. The pellet was washed with 1.5 mL ice cold 75% ethanol. Once the ethanol was removed, the pellet was allowed to dry for 15-20 minutes at room temperature before being resuspended in sterile water.

The RNA concentration was measured in NanoDrop and 1 µg RNA was run on a 0.8% agarose gel to assess RNA quality.

cDNA synthesis

2 µg of total RNA was treated with DNase I (Thermo Scientific) for 30 minutes at 37°C (2 µg RNA, 0.1 U/µL DNase I, 1X DNase buffer) followed by 10 minutes at 65°C for DNase inactivation. For one set of replicates, the DNase inactivation was performed in the presence of 5 mM EDTA and for the other there was no EDTA. Reverse transcription was performed using RevertAid H Minus Reverse Transcriptase (Thermo Scientific) with 10 minutes incubation at 25°C followed by 1 hour at 42°C and 10 minute inactivation at 70°C (10 U/µL reverse transcriptase, 1X reaction buffer, 1 mM dNTPs, 3.75 µM oligo dT primer, 1.25 µM random hexamer primer, 0.5 U/µL RiboLock RNase inhibitor). The cDNA was diluted 5 times before using it for RT-qPCR.

RT-qPCR

RT-qPCR was performed using 2X Maxima SYBR Green/ROX qPCR Master Mix (ThermoFisher Scientific) according to manufacturer's instructions. The reaction was performed in StepOnePlus™ Real-Time PCR System (Applied Biosystems) with an initial denaturation at 95°C for 10 minutes followed by 40 cycles of 95°C for 15 seconds and 1 minute at 60°C for annealing and elongation, followed by melt curve analysis. The control

gene used was *eif3c* and three ERGO-1 target genes were measured, C40A11.10, E01G4.5 and E01G4.7. Primer sequences for each of the genes can be found in Table S1.

Gene expression for each sample was quantified as fold change using the ΔΔCt method, normalizing target gene expression to the expression of *eif3c*. ΔCt was calculated by subtracting the cycle threshold (Ct) value of *eif3c* from the Ct value of the target gene for each sample, and ΔΔCt was calculated by subtracting the ΔCt value of the wild type control sample from the ΔCt value of each sample. Fold change was then calculated as $2^{-\Delta\Delta Ct}$.

Crosslinking and protein extraction

100,000 bleach synchronized worms were grown up on 8P-enriched peptone plates with NA22 and collected as young gravid adults in 1X M9. The worms were washed 3 times with 1X M9 and resuspended in M9 following the washes. The worms in M9 were rocked for 20 minutes to allow worms to digest any bacteria present in their intestine. The worms were then collected and resuspended in M9 buffer and plated on clean NGM plates, making sure the worms were in one layer and avoiding areas of high worm concentration. The NGM plates were placed into a 254 nm UV crosslinker and irradiated with 500 mJoule/cm^2. The worms were collected in a microcentrifuge tube and resuspended in ice-cold homogenization buffer (100 mM NaCl, 25 mM HEPES, 250 µM EDTA, 0.1% TX-100, 2 mM DTT, 25 U/mL RiboLock RNase inhibitor, HALT Protease Inhibitor) and sonicated 5 times with 10 second pulses and 50 seconds resting on ice in between pulse, using the 50% intensity setting on a Sonics Vibra-Cell™ model VCX130 sonicator (Sonics & Materials Inc.). The lysates were centrifuged at 16,000 g for 15 minutes at 4°C and the supernatants were transferred to a new microcentrifuge tube. The lysates were frozen in liquid nitrogen and stored at -80°C.

A project plan for CLIP was created in conjunction with Erik Holmqvist's group using a modified EasyCLIP protocol (Porter *et al.* 2021), where fluorescently labeled adapters are used to identify the RNA-protein complex in SDS-PAGE (Figure 11). In the modified protocol, the crosslinked protein samples would be immunoprecipitated using anti-FLAG beads and fluorescent 3' and 5' adapters then ligated to the RNA, with the 3' adaptor containing an oligo(A) sequence. Following the ligation, the samples would be eluted, separated by SDS-PAGE and transferred to a membrane. Fluorescence imaging would then be used to identify the location of the RBP-RNA complex on the membrane and the area around the signal cut out of the membrane. RNA would then be extracted from the membrane

and purified with oligo dT bead purification, and used for cDNA synthesis followed by sequencing (Figure 2).

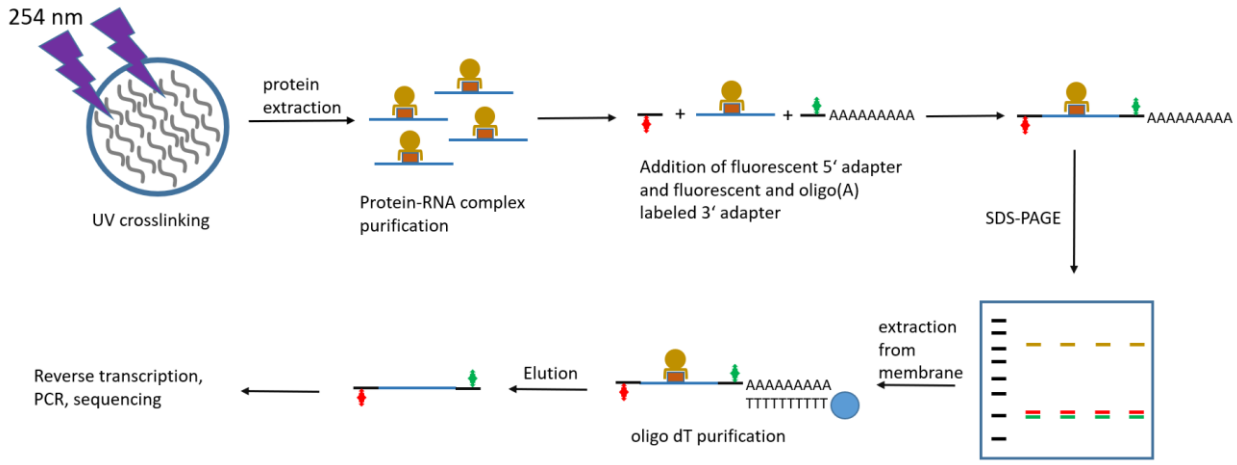

Figure 2. Overview of the workflow for easyCLIP-sequencing.

Immunofluorescence staining

N2 and AHS205 hermaphrodite worms were stained with primary antibodies Monoclonal ANTI-FLAG® M2 antibody produced in mouse (Sigma-Aldrich) and secondary antibodies Alexa Fluor™ 488 F(ab')2 from goat anti-mouse IgG (Thermo Fisher). Young adult worms were picked into a drop of 10 mM levamisole in M9 on a poly-L-lysine slide and dissected with a scalpel to extrude the gonads. A cover slip was then placed on top of the drop and excess liquid was aspirated until worms were slightly compressed. The slides were transferred into liquid nitrogen for a minimum of 10 minutes for freeze-cracking. Once the slides were removed from the liquid nitrogen, the cover slip was flicked off and the slide was immediately transferred into a humid chamber and 4% formaldehyde in 1X PBS was added to the slide for fixation. The slides were incubated for 30 minutes at 4°C and then rinsed in PBST (1X PBS with 0.1% Tween). Primary antibody diluted 1:1000 in 1% non-fat milk was then added to the slides and incubated for 1 hour at room temperature. The slides were washed 3 times in PBST before the secondary antibody diluted 1:1000 in 1% non-fat milk was added to the slides and incubated in the dark for 1 hour at room temperature. The slides were washed 3 times in PBST before the slides were mounted in VectaShield mounting medium with DAPI and sealed with nail polish. The slides were then imaged with a Ecplise Ti fluorescence microscope (Nikon) and Zyla sCMOS camera (Andor Technology).

Results

Outcrossing of C. elegans mutations affecting different fubl-1 isoforms

Prior RNA sequencing analysis of data from Legnini *et al.* (2019) has indicated that the vast majority of *fubl-1* transcripts in *C. elegans* match the *fubl-1a* isoform, with a small amount of transcripts matching the *fubl-1c* isoform and even fewer transcripts matching the *fubl-1b* isoform (Pålsson 2022). However the contribution of the different isoforms to the function of FUBL-1 has not yet been determined. One main objective of this study was therefore to investigate the roles of different FUBL-1 isoforms in the 26G RNA pathway.

At the start of this project, Hinas lab had *C. elegans* strains with a *fubl-1* deletion likely causing *fubl-1* knockout, as well as strains with single nucleotide substitution causing premature stop codons in the different isoforms. AHS158 contains a *fubl-1* knockout deletion (*fubl*-1(tm2769)V), AHS170 contains a premature stop codon affecting isoform a (*fubl-1*(gk660081)V), VC40712 contains a premature stop codon affecting isoform b (*fubl-1*(gk773341)V), and AHS171 contains a premature stop codon affecting all three isoforms a, b and c (*fubl-1*(gk668481)V) (Figure 3). The mutant strains were created through random mutagenesis and therefore may have other mutations elsewhere in the genome; to create worms with the mutations of interest but otherwise near-wild type genomes, outcrossing had to be performed. All the aforementioned strains were outcrossed 6x before experiments could be done on them.

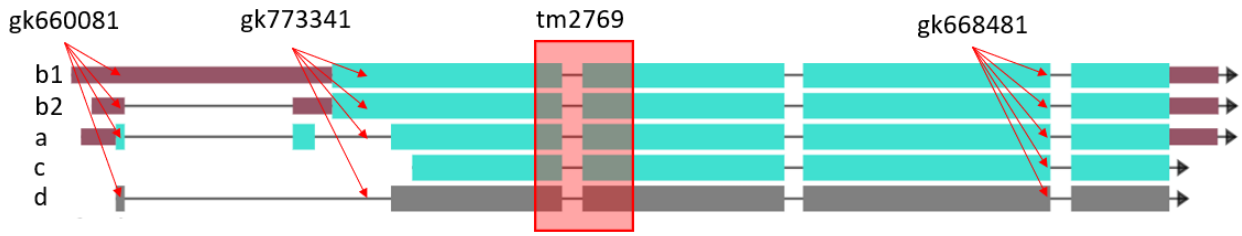

Figure 3. Schematic of the fubl-1 *isoforms with arrows pointing to the location of each point mutation and a box indicating the area of the deletion mutation used in this project.*

Wild type males (N2 strain) were first obtained through heat shock and put on plates with L4 stage hermaphrodites of the mutant strain to be outcrossed. After 4 days, L4 stage males were picked from the cross plates and mated to N2 L4 stage hermaphrodites. After 4 more days, L4 stage hermaphrodites were picked from the new cross plates and placed individually onto small plates for reproduction by self-fertilization followed by genotyping. The following day, the adult hermaphrodite was picked from each plate and used for genotyping PCR to identify

heterozygotes. From heterozygote plates, progeny were picked onto individual plates and once they had laid embryos, single worm PCR was performed again to identify homozygotes (Figure 4).

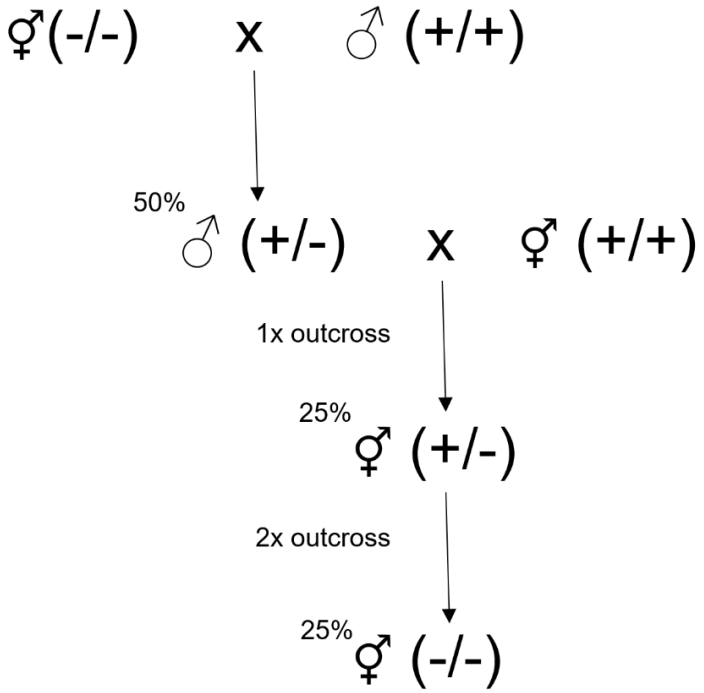

Figure 4. Schematic showing how outcrossing of autosomal genes such as fubl-1 *is performed; each outcrossing round provides two times outcrossing and the scheme is repeated until 6 times outcrossing is reached.*

Single worm PCR of the *fubl-1* fragment containing each mutation was used along with restriction digests for genotyping. The *fubl-1a* isoform mutation (gk660081) was identified by digestion with BveI (Figure 5A) and gel electrophoresis (Figure 6A), the *fubl-1b* isoform mutation (gk773341) by digestion with TasI (Figure 5B) and gel electrophoresis (Figure 6B), and the mutation affecting all isoforms (gk668481) was identified by digestion with BseGI (Figure 5C) and gel electrophoresis (Figure 6C). No enzyme digestion was needed to identify the *fubl-1* deletion mutants (tm2769) as the deletion was large enough to cause a visible size difference with agarose gel electrophoresis of PCR-amplified fragments containing the deleted region; the wild type and mutant products were 538 and 314 bp respectively (Figure 6D).

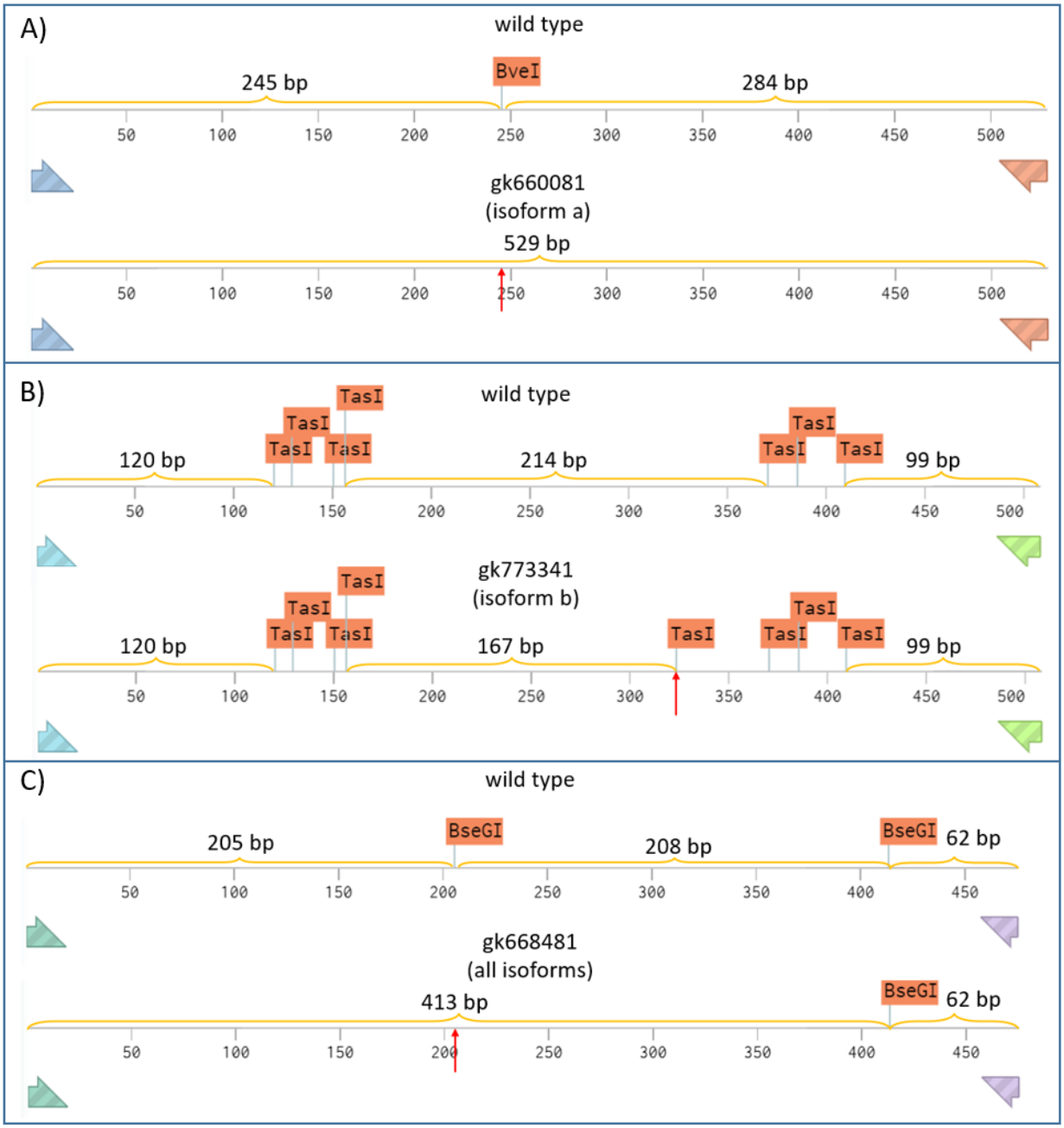

Figure 5. Schematic of the PCR fragments covering the different fubl-1 *isoform mutations, with restriction sites used for strain identification as well as the size of the expected bands for restriction digest indicated. The red arrows point to the site of the point mutation for each mutant.*

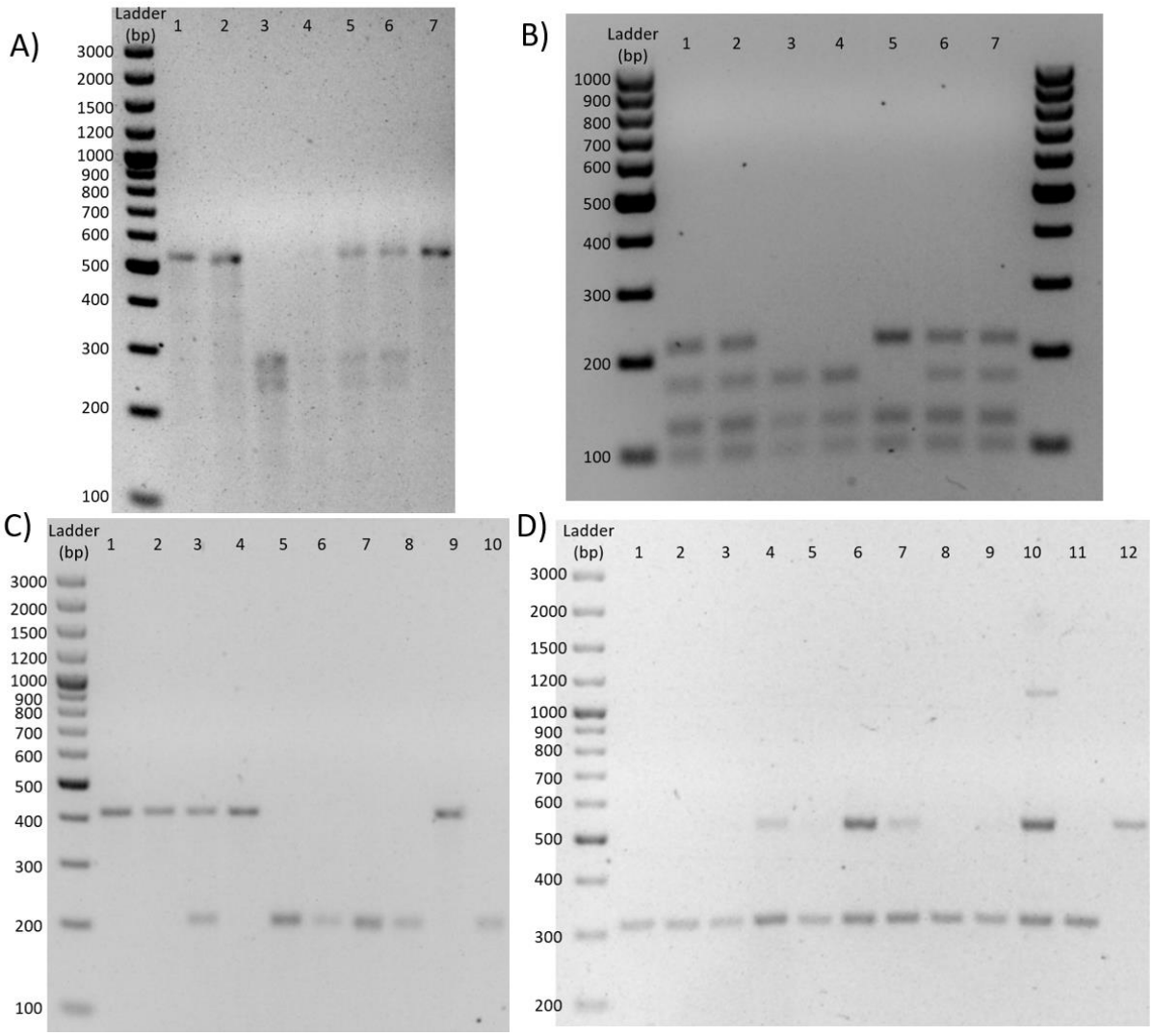

Figure 6. Gel electrophoresis results for genotyping of different fubl-1 *mutants by single worm PCR and restriction digest. A) PCR products from* fubl-1a *(gk660081) mutant outcrossing digested with BveI. Lane 3 is homozygous mutant, lanes 5 and 6 are heterozygous and lanes 1, 2 and 7 are homozygous wild type. B) PCR products from* fubl-1b *(gk773341) mutant outcrossing digested with TasI. Lanes 1, 2, 6 and 7 are heterozygous, lanes 3 and 4 are homozygous mutants and lane 5 is homozygous wild type. C) PCR products from outcrossing of worms with a* fubl-1 *mutation affecting all isoforms (gk668481) digested with BseGI. Lanes 1, 2, 4 and 9 are homozygous mutants, lane 3 is heterozygous and 5, 6, 7, 8 and 10 are homozygous wild type. D) PCR products from outcrossing of worms with a* fubl-1 *deletion mutation (tm2769). Lanes 1, 2, 3, 5, 8, 9, 11 are homozygous mutants, lanes 4, 6, 7 and 10 are heterozygous and lane 12 is homozygous wild type.*

RT-qPCR indicates partial redundancy of fubl-1 *isoforms*

Prior research conducted by the Hinas group has shown an upregulation of ERGO-1 target genes C40A11.10, E01G4.5 and E01G4.7 in *fubl-1* deletion mutants compared to wild type worms (Roy *et al.*, manuscript in preparation). The expression levels of said target genes have not yet been investigated in the various *fubl-1* isoform nonsense mutants and this study aimed to do that to provide further insight into the role of the different *fubl-1* isoforms in the ERGO-1 pathway.

Total RNA was extracted from the different mutant strains and RNA quality was assessed with gel electrophoresis. In all cases there were bands at approximately 1 kb and above 1.5 kb representing 18S and 28S rRNA respectively (Figure 7), indicating there was not major RNA degradation occurring.

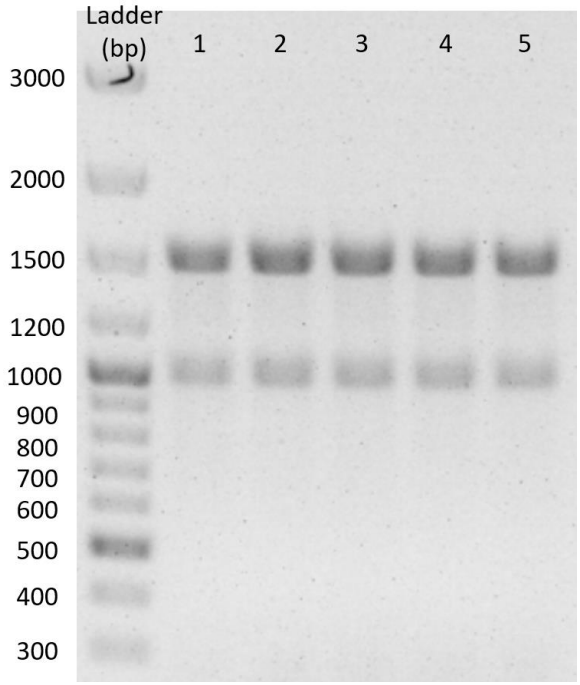

Figure 7. Gel electrophoresis of total RNA extracted from biological replicate 2.

RT-qPCR was performed for the ERGO-1 target genes listed above, using *eif3c* as a control gene. For one biological replicate, RT-qPCR for *eif3c* showed an acceptable cycle threshold value for all the strains and low standard deviation between technical replicates for all but isoform b mutants. For all three of the ERGO-1 target genes, the *fubl-1* deletion mutant showed the highest increase in expression compared to the isoform mutants. Relative expression of C40A11.10 showed upregulation in all of the isoform mutants as well as the deletion mutant, with the isoform a mutant showing expression closest to the deletion mutant

(Figure 8A). For E01G4.7, the isoform mutants showed a decrease in expression (Figure 8B), and for E01G4.5 isoform a and all isoform mutants showed a slight increase in expression while the isoform b mutant showed a slight decrease in expression (Figure 8C). The *eif3c* control gene qPCR for the isoform b mutant showed a higher standard deviation for technical replicates than the other samples, indicating that results for this strain are less reliable.

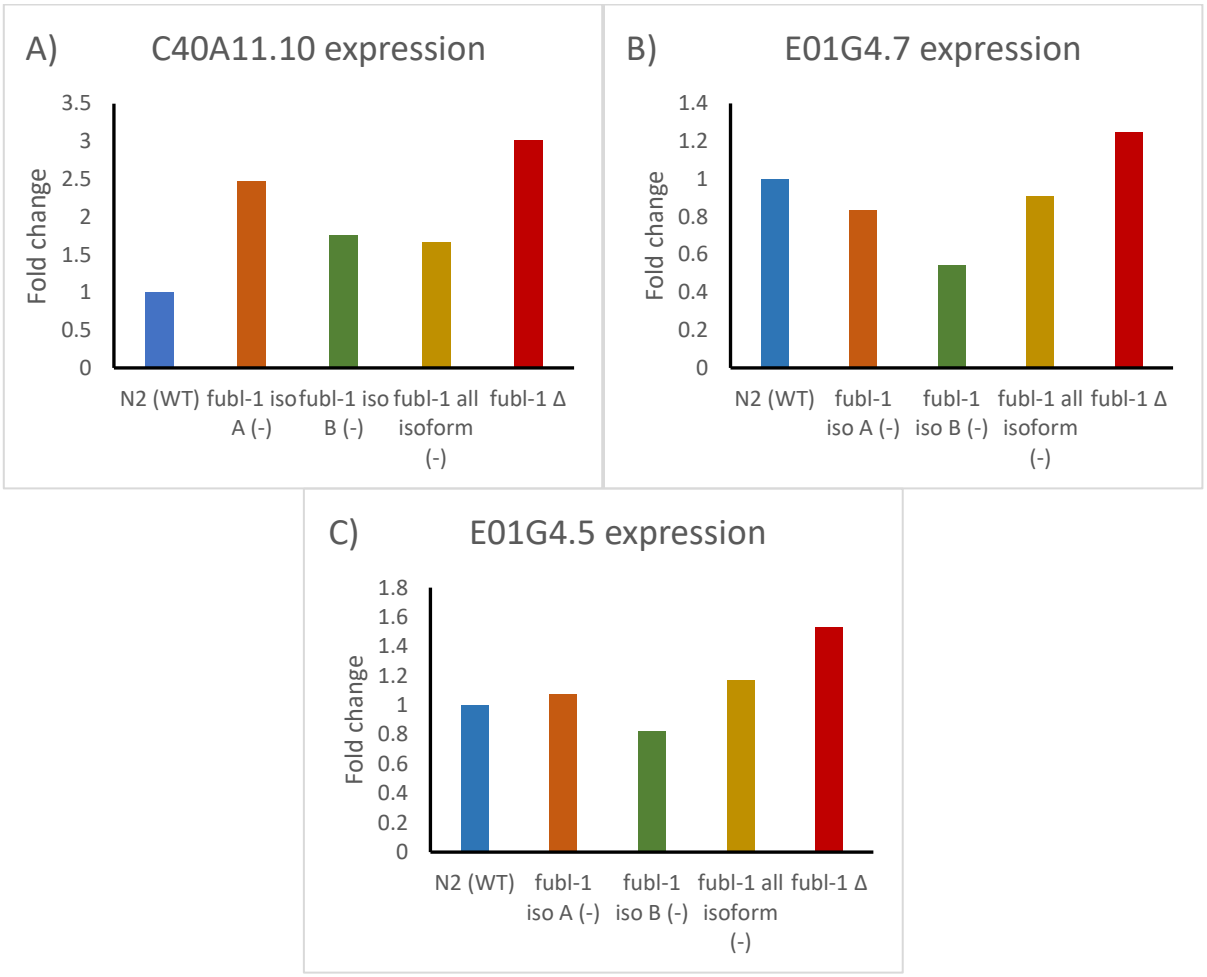

Figure 8. A) Relative expression levels of C40A11.10 in five different strains. B) Relative expression levels of E01G4.7 in five different strains. C) Relative expression levels of E01G4.5 in five different strains.

For subsequent biological replicates, qPCR results for *eif3c* were either undetectable or showed a high standard deviation. For one attempt, there was no EDTA added during the inactivation of the DNase prior to reverse transcription. EDTA acts as a chelator which stabilizes the RNA and prevents its hydrolysis during the DNase inactivation step, and therefore one would expect a high qPCR cycle threshold due to RNA degradation prior to reverse transcription. Despite this, qPCR results showed a very low cycle threshold for the control gene *eif3c* and indicated too high cDNA concentration. After a dilution series was

performed, the cycle threshold value was within acceptable range, but the standard deviation between technical replicates was very high and the data could therefore not be used (data not shown). Reverse transcription was repeated for the second biological replicate and performed twice on the third biological replicate with EDTA added during DNase inactivation, but control gene qPCR showed little to no detected cDNA. When qPCRs were performed using the same reagent mix on the same plate using cDNA from the first biological replicate the results for *eif3c* showed an acceptable cycle threshold value and low standard deviation between technical replicates, and therefore the problem seems to lie in the reverse transcription step rather than the qPCR step. Further troubleshooting was not possible due to time constraints and as a result there is only RT-qPCR data available for one biological replicate.

Immunofluorescence staining confirms nuclear expression of FUBL-1 in the germline

While FUBL-1 appears to play a role in the ERGO-1 26G RNAi pathway which is active in the oogenic germline and embryos, it has not yet been confirmed whether FUBL-1 is expressed in the germline. Prior research has been conducted by the Hinas group looking at FUBL-1 localization using a *fubl-1::mCherry* transgene, finding broad expression of FUBL-1 in various somatic tissue, including cytoplasmic expression in the pharynx, but little or no expression in the germline. However as the *fubl-1::mCherry* was an extrachromosomal array and there is general repression of repetitive transgenic DNA in the *C. elegans* germline (Kelly & Fire 1998), the *fubl-1::mCherry* transgene was not appropriate for assessment of germline expression. In this study I have used a *C. elegans* strain with 3xFLAG tag inserted into the genomic *fubl-1* locus and therefore *fubl-1::3xFLAG* should not be subject to transgene silencing in the germline.

To assess localization of FUBL-1, adult AHS205 (*fubl-1::3xFLAG*) and N2 (wild type) hermaphrodites were stained with anti-FLAG antibodies and mounted in mounting medium with DAPI. The worms were then imaged using a wide-field fluorescence microscope. The staining showed nuclear expression of FUBL-1 in the gonad (Figure 9)

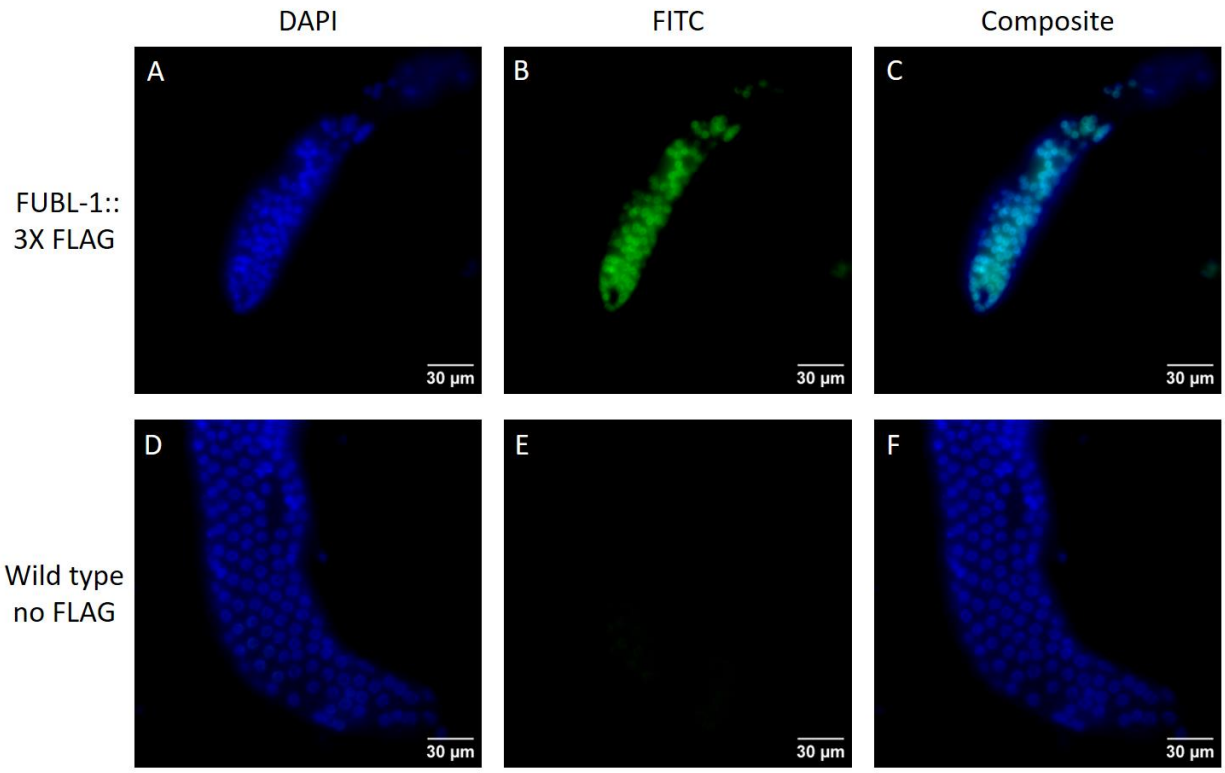

Figure 9. Wide-field fluorescence microscope images of adult hermaphrodite gonads from FUBL-1::3xFLAG (A-C) and wild type (D-F) worms stained with anti-FLAG antibody and DAPI.

While in the FUBL-1::mCherry imaging study done previously showed expression of FUBL-1 was broad and cytoplasmic in the pharynx (Figure 10), the FUBL-1::3xFLAG staining showed some nuclear expression of FUBL-1 (Figure 11).

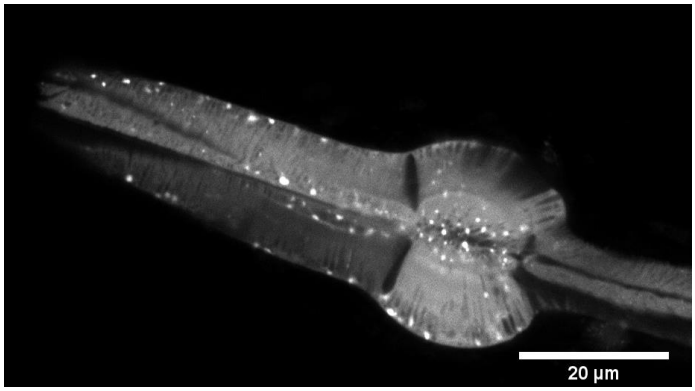

Figure 10. FUBL-1::mCherry expression in the pharynx (Roy et al., manuscript in preparation).

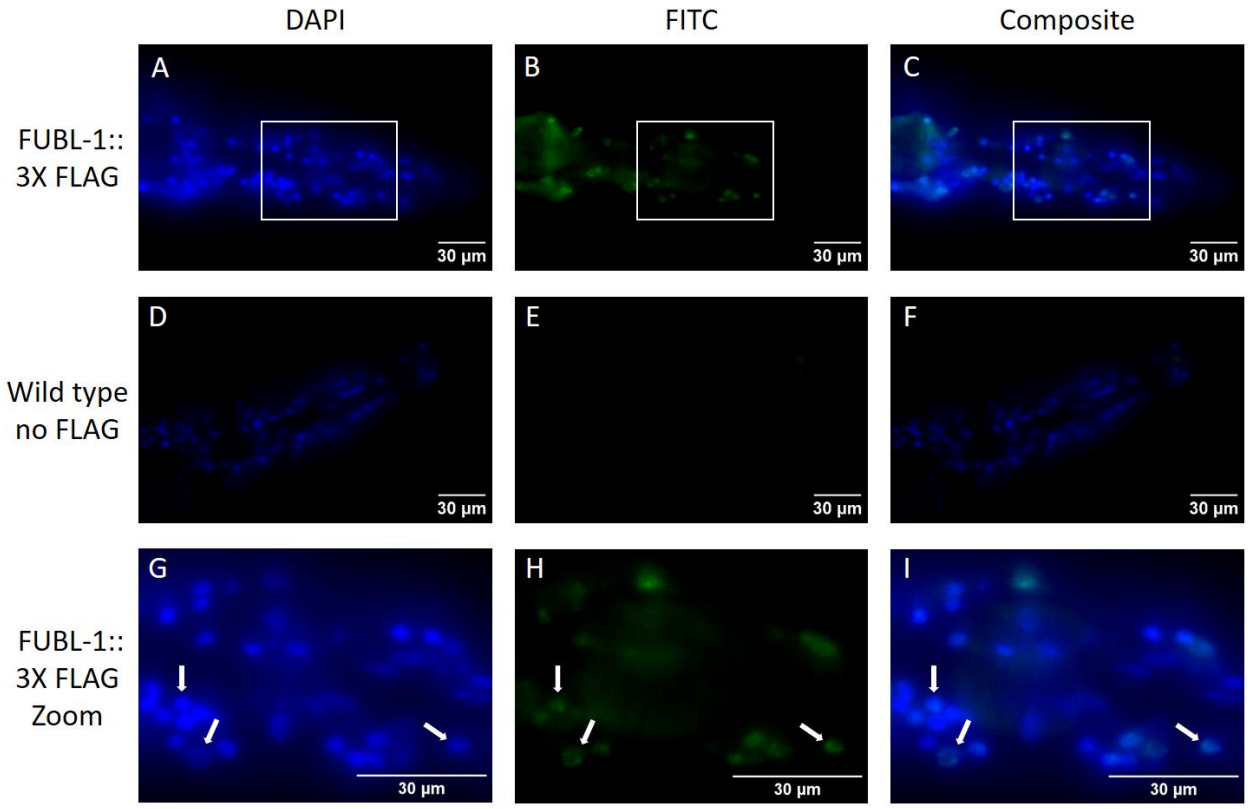

Figure 11. Wide-field fluorescence images of adult hermaphrodite pharynx anti-FLAG antibodies and DAPI. A-C) FUBL-1::3xFLAG worms stained with anti-FLAG and DAPI. D-F) Wild type worms stained with anti-FLAG and DAPI. G-I) A portion of A-C zoomed in, with arrows pointing to nuclei showing FUBL-1 expression.

CLIP sequencing to identify potential FUBL-1 target RNAs

As FUBL-1 is predicted to be an RNA binding protein, identifying potential RNA targets of FUBL-1 would help to provide information about its function; one way of doing this is through crosslinking and immunoprecipitation sequencing (CLIP-seq). As part of this study, a project plan for CLIP sequencing was created using a modified EasyCLIP protocol (Porter *et al.* 2021).

To initiate this study, growth of large amounts of worms was optimized. The protocol for UV crosslinking was set up and applied to 100,000 adult hermaphrodite worms, followed by protein extraction. Unfortunately, the remainder of the CLIP experiment could not be performed yet due to problems with adapter ligation and time constraints.

Discussion

There are many questions to be answered as to the function of FUBL-1. FUBL-1 is an RNA-binding protein with a predicted three functional isoforms (a, b and c) (WormBase website, https://wormbase.org/species/c_elegans/gene/WBGene00007534, release WS288, 2023-05-10), however the function of the different isoforms is yet to be confirmed. Prior research by the Hinas group has found an increase in exo-RNAi in FUBL-1 deletion mutants (Roy *et al.*, manuscript in preparation), comparable to the Eri phenotype seen in ERGO-1 deficient worms (Lee *et al.* 2006). They have also found that *fubl-1* deletion mutants show an upregulation of ERGO-1 target genes, further indicating a role of FUBL-1 in the ERGO-1 pathway (Roy *et al.*, manuscript in preparation).

In this study, qPCR of ERGO-1 targets in *C. elegans* strains with *fubl-1* deletion as well as strains with premature stop codons affecting isoform a, isoform b, or isoform a, b and c shows a potential upregulation of ERGO-1 targets consistent with prior results. The upregulation in isoform mutants appeared lower than in the deletion mutants, which may indicate partial redundancy of the different isoforms. The worms with the gk668481 mutation, which results in a premature stop codon affecting isoforms a, b and c, also shows less upregulation of ERGO-1 target genes than the deletion mutants. This is unexpected as all functional isoforms should be inactivated in this mutant. If this holds true for further biological replicates, it may indicate that isoform d is functional and partially redundant with the other isoforms. While the open reading frame for isoform d is only 20 amino acids, this does not preclude functionality as there are instances of short polypeptides with lengths under 100 amino acids having a functional role (Su *et al.* 2013). Alternatively, the gk668481 mutation is located towards the end of the mRNA for all of the isoforms, and it could therefore be possible that a truncated version of the FUBL-1 protein retains some function. Due to the fact that this is only one biological replicate, no strong conclusions can be drawn. More qPCR experiments should be done in the future to see if the findings hold true with other biological replicates. Repeating the experiments using mutants with a premature stop codon affecting isoforms a, b and c earlier in the coding sequence could be done to investigate whether the truncated protein retains partial function.

As the ERGO-1 pathway is active in the germline (Han *et al.* 2009), knowing whether FUBL-1 is expressed in the germline will help to confirm the connection of FUBL-1 to the ERGO-1 pathway. FUBL-1 expression has previously been examined using a FUBL-1::mCherry

fusion transgene on an extrachromosomal array (Roy *et al.*, manuscript in preparation). However, as this transgene was overexpressed, expression patterns may differ from endogenous FUBL-1 expression patterns as well as be affected by transgene silencing, particularly in the germline. In this study I have utilized worms with *fubl-1::3xFLAG* which is inserted into the genomic *fubl-1* locus and should therefore show endogenous FUBL-1 expression patterns and levels. Immunofluorescence staining confirmed that FUBL-1 is expressed in the germline nucleus as well as showing that there is some nuclear expression in the pharynx. As the imaging was done with a wide-field microscope, in the future it would be helpful to repeat the staining to get better images using a confocal microscope. Adding a blocking step to the staining process could also help to get clearer images by reducing background fluorescence. The anti-FLAG staining turned out to be specific and successful in imaging FUBL-1 localization and it would be helpful in the future to perform more staining experiments to assess FUBL-1 expression in other tissues. As the Eri phenotype in ERGO-1 mutants shows maternal rescue (Zhuang & Hunter 2011, Almeida *et al.* 2019b), it would be interesting to see whether FUBL-1 is maternally provided to progeny. This could be done by staining embryos from heterozygous *fubl-1::3xFLAG/+* worms.

While FUBL-1 is predicted to be an RNA-binding protein, its RNA targets have yet to be identified. Using the protein extracted from crosslinked worms in this study following the CLIP workflow explained earlier will give insight into the targets of FUBL-1 and the function of the protein.

Acknowledgements

I would like to thank Andrea Hinas for her guidance in creating this thesis and giving me the opportunity to do my project in her lab. I would like to thank Erik Holmqvist's lab, particularly Kim Boi Le Huyen, for their help with the CLIP portion of this project. Finally, I would like to thank the other members of the Microbiology and Immunology section of the Cell and Molecular Biology department at Uppsala University for offering assistance and guidance at various different points in my project, as well as WormBase.

References

Almeida MV, Andrade-Navarro MA, Ketting RF. 2019a. Function and Evolution of Nematode RNAi Pathways. Non-Coding RNA 5: 8.

Almeida MV, de Jesus Domingues AM, Ketting RF. 2019b. Maternal and zygotic gene regulatory effects of endogenous RNAi pathways. PLoS Genetics 15: e1007784.

Brenner S. 1974. The Genetics of CAENORHABDITIS ELEGANS. Genetics 77: 71–94.

Buckley B, Burkhart K, Gu SG, Spracklin G, Kershner A, Fritz H, Kimble J, Fire A, Kennedy S. 2012. A nuclear Argonaute promotes multi-generational epigenetic inheritance and germline immortality. Nature 489: 447–451.

Burton NO, Burkhart KB, Kennedy S. 2011. Nuclear RNAi maintains heritable gene silencing in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 108: 19683–19688.

Evans T. 2006. Transformation and microinjection. WormBook, doi 10.1895/wormbook.1.108.1.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391: 806–811.

Guang S, Bochner AF, Burkhart KB, Burton N, Pavelec DM, Kennedy S. 2010. Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. Nature 465: 1097–1101.

Guang S, Bochner AF, Pavelec DM, Burkhart KB, Harding S, Lachowiec J, Kennedy S. 2008. An Argonaute transports siRNAs from the cytoplasm to the nucleus. Science (New York, NY) 321: 537–541.

Han T, Manoharan AP, Harkins TT, Bouffard P, Fitzpatrick C, Chu DS, Thierry-Mieg D, Thierry-Mieg J, Kim JK. 2009. 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 106: 18674–18679.

Hannon GJ. 2002. RNA interference. Nature 418: 244–251.

Kelly WG, Fire A. 1998. Chromatin silencing and the maintenance of a functional germline in Caenorhabditis elegans. Development (Cambridge, England) 125: 2451–2456.

Kim JK, Gabel HW, Kamath RS, Tewari M, Pasquinelli A, Rual J-F, Kennedy S, Dybbs M, Bertin N, Kaplan JM, Vidal M, Ruvkun G. 2005. Functional Genomic Analysis of RNA Interference in C. elegans. Science 308: 1164–1167.

Lacal I, Ventura R. 2018. Epigenetic Inheritance: Concepts, Mechanisms and Perspectives. Frontiers in Molecular Neuroscience 11:

Lee RC, Hammell CM, Ambros V. 2006. Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. RNA 12: 589–597.

Legnini I, Alles J, Karaiskos N, Ayoub S, Rajewsky N. 2019. FLAM-seq: full-length mRNA sequencing reveals principles of poly(A) tail length control. Nature Methods 16: 879–886.

Li S, Armstrong CM, Bertin N, Ge H, Milstein S, Boxem M, Vidalain P-O, Han J-DJ, Chesneau A, Hao T, Goldberg DS, Li N, Martinez M, Rual J-F, Lamesch P, Xu L, Tewari M, Wong SL, Zhang LV, Berriz GF, Jacotot L, Vaglio P, Reboul J, Hirozane-Kishikawa T, Li Q, Gabel HW, Elewa A, Baumgartner B, Rose DJ, Yu H, Bosak S, Sequerra R, Fraser A, Mango SE, Saxton WM, Strome S, van den Heuvel S, Piano F, Vandenhaute J, Sardet C, Gerstein M, Doucette-Stamm L, Gunsalus KC, Wade Harper J, Cusick ME, Roth FP, Hill DE, Vidal M. 2004. A Map of the Interactome Network of the Metazoan C. elegans. Science (New York, NY) 303: 540–543.

Pålsson J. 2022. Differential expression and function of fubl-1 gene isoforms in C. elegans. B.Sc. degree project, Uppsala University

Porter DF, Miao W, Yang X, Goda GA, Ji AL, Donohue LKH, Aleman MM, Dominguez D, Khavari PA. 2021. easyCLIP analysis of RNA-protein interactions incorporating absolute quantification. Nature Communications 12: 1569.

Rankin CH. 2015. A review of transgenerational epigenetics for RNAi, longevity, germline maintenance and olfactory imprinting in Caenorhabditis elegans. Journal of Experimental Biology 218: 41–49.

Shiu PK, Hunter CP. 2017. Early developmental exposure to dsRNA is critical for initiating efficient nuclear RNAi in C. elegans. Cell reports 18: 2969–2978.

Su M, Ling Y, Yu J, Wu J, Xiao J. 2013. Small proteins: untapped area of potential biological importance. Frontiers in Genetics 4:

Tabach Y, Billi AC, Hayes GD, Newman MA, Zuk O, Gabel H, Kamath R, Yacoby K, Chapman B, Garcia SM, Borowsky M, Kim JK, Ruvkun G. 2013. Small RNA pathway genes identified by patterns of phylogenetic conservation and divergence. Nature 493: 694–698.

Vastenhouw NL, Brunschwig K, Okihara KL, Müller F, Tijsterman M, Plasterk RHA. 2006. Long-term gene silencing by RNAi. Nature 442: 882–882.

WormBase website. fubl-1 (gene) - WormBase : Nematode Information Resource. online: https://wormbase.org/species/c_elegans/gene/WBGene00007534#0-9f1356e-10-1356. Accessed May 10, 2023.

Wu E, Vashisht AA, Chapat C, Flamand MN, Cohen E, Sarov M, Tabach Y, Sonenberg N, Wohlschlegel J, Duchaine TF. 2017. A continuum of mRNP complexes in embryonic microRNA-mediated silencing. Nucleic Acids Research 45: 2081–2098.

Zhuang JJ, Hunter CP. 2011. Tissue Specificity of Caenorhabditis elegans Enhanced RNA Interference Mutants. Genetics 188: 235–237.

Supplement

Table S1. List of primers used for PCR and qPCR.

Primer name	Primer sequence	Target
C12D8.1_del_for	atcttagttcctgtttccatgg	Amplification of FUBL-1 around tm2769 deletion
C12D8.1_del_rev	ctttagaaggatcccgttcc	
C12D8.1_gk773341_for	gttgataagatgtttcgcgg	Amplification of FUBL-1 around gk773341 point mutation
C12D8.1_gk773341_rev	ctttgttcagcaagctactcac	
C12D8.1_gk660081_for	gggaatcataaattcaaattgc	Amplification of FUBL-1 around gk660081 point mutation
C12D8.1_gk660081_rev	gtgtttttttacttgtgtaggcttg	
C12D8.1_gk668481_for	ctgtgctgacatttgaggc	Amplification of FUBL-1 around gk668481 point mutation
C12D8.1_gk668481_rev	caatggaatccggtagca	
301_eif-3.c_fwd	acacttgacgagcccaccgac	Amplification of reference gene *eif3c*
302_eif-3.c_rev	tgccgctcgttccttcctgg	
317_C40A11.10 F	tgtggatttcaacgtggcgg	Amplification of ERGO-1 target C40A11.10
318_C40A11.10 R	gatgctatcgcttagcggtgg	
315_E01G4.5 F 362_E01G4.5_Rev3	ctcaagaaagtttcacagcaggcc cacaatcacggcacacaaaac	Amplification of ERGO-1 target E01G4.5
307_E01G4.7 F	gcacaaggtttcgttcttggtg	Amplification of ERGO-1 target E01G4.7
308_E01G4.7 R	agtgacatcccttctgatcg	