

Regulation of the central apoptotic pathway during *C. elegans* development

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften

Doctor rerum naturalium (Dr. rer. nat.) an der Fakultät für Biologie der

Ludwig-Maximilians-Universität München



Sebastian Lühr

Haan, 2021

Diese Dissertation wurde angefertigt
unter der Leitung von Prof. Dr. Barbara Conradt
im Bereich der Zellbiologie und Entwicklungsbiologie
an der Ludwig-Maximilians-Universität München

1. Gutachter: Prof. Dr. Barbara Conradt

2. Gutachter: Prof. Dr. Heinrich Leonhardt

Tag der Einreichung: 24.11.2020

Tag der mündlichen Prüfung: 22.04.2021

Eidesstattliche Versicherung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist. Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist. Ich habe mich nicht anderweitig einer Doktorprüfung ohne Erfolg unterzogen.

Sebastian Lühr (Haan, 26. April 2021)

Publications originating from this thesis

Sherrard R, Luehr S, Holzkamp H, McJunkin K, Memar N and Conradt B.
miRNAs cooperate in apoptosis regulation during *C. elegans* development.
Genes Dev. 2017 Jan 15;31(2):209-222. doi: 10.1101/gad.288555.116. Epub 2017 Feb 6.

Acknowledgements

I would like to express my sincere gratitude and appreciation to my supervisor, Prof. Dr. Barbara Conradt, for the excellent supervision, the immense scientific input during my research, and for always having an open door. Further I want to thank you for the trust and support towards my other passion, IT, by letting me handle the workgroup servers, computers and writing scripts for data analysis. Also, the Coursera course has been a special experience I enjoyed, and I am still proud of.

I would like to thank Dr. Nadin Memar for her outstanding support during my laboratory work and teaching me all the techniques; for always having an open door, also for supporting me through personal ups and downs and always motivating me in times of frustration. It was a pleasure to work with you.

I would also like to thank everyone I have had the pleasure to work with: Jeffrey, Eric, Ryan, Stephane, Marion, Kyoko, Heinke, Laura, Anna, Jennifer, Elisabeth, Nadja, Melanie, Linda, Fabian, Fabio, Simon, Sriyash, Esther, Tamara, Hai, Bo, Saroj, Sayantan and Tobias. You were always keen to push the boundaries of science forward, sharing your knowledge and practical experience, creating an awesome atmosphere. I really enjoyed being part of the team.

I would also like to thank my parents, for making my studies in biology possible; for their support and the endless help and motivation in every situation at every time, day or night. I am grateful to have you in my life.

Finally, I like to thank my girlfriend Tanja for having endless patience with me during my studies. For always keeping me happy and motivated and for pushing me when it was needed. I am happy to have you on my side.

I also want to thank WormBase for the excellent community platform. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440)

Contents

Eidesstattliche Versicherung	i
Publications originating from this thesis	ii
Acknowledgements	iii
Contents	iv
Informative abstract	1
1 Introduction	3
1.1 <i>Caenorhabditis elegans</i> – a model organism	4
1.2 Programmed cell death.....	6
1.2.1 Apoptosis.....	6
1.2.2 Apoptosis in <i>Caenorhabditis elegans</i> development.....	7
1.2.3 The central cell death pathway in <i>Caenorhabditis elegans</i>	9
1.2.4 EGL-1 key integrator of apoptotic stimuli	12
1.3 Chromatin structure, remodeling influence gene expression	14
1.3.1 Chromatin structure	14
1.3.2 Telomeres protect the end of linear chromosomes	15
1.3.3 Chromatin modifications	16
1.4 Post-transcriptional regulation of gene expression.....	17
1.4.1 Mechanisms of post-transcriptional regulation	17
1.4.2 Post-transcriptional regulation via the 3' untranslated region	19
1.5 Post-transcriptional regulation by MicroRNAs	21
1.5.1 MicroRNAs, biogenesis and function	21
1.5.2 miRNA-mediated regulation of programmed cell death.....	23
1.5.3 miR-35-family and miR-58-family miRNAs in <i>C. elegans</i>	24

1.6	CRISPR/Cas system	26
1.6.1	Targeted genome modification	27
1.6.2	Visualizing distinct DNA loci.....	28
1.7	Aim of this study	30
2	Results	32
2.1	Preface	33
2.2	Post-transcriptional regulation of <i>egl-1</i> by microRNA	34
2.2.1	The 3' untranslated region of <i>egl-1</i> contains conserved binding sites for <i>microRNAs</i>	34
2.2.2	Abnormally large-cell corpses are present in <i>mir-35</i> family mutants.....	36
2.2.3	Abnormally large-cell corpse phenotype in miR-35 family mutants is enhanced by additional loss of miR-58 family miRNAs	41
2.2.4	The abnormally large-cell death phenotype is dependent on the central apoptotic machinery	42
2.2.5	Downregulation of <i>egl-1</i> mRNA expression is mediated by the 3'UTR.....	44
2.2.6	Mothers and sisters of normal apoptotic cells die inappropriately in <i>mir-35</i> family mutants	47
2.3	Analysis of the CRISPR generated <i>egl-1</i> 3'UTR variant <i>bc275</i>	49
2.3.1	Generation of <i>egl-1(bc275)</i> via CRIPR/Cas9 co-conversion	49
2.3.2	The gene variant <i>egl-1 (bc275)</i> is causing a general block in apoptosis.....	52
2.3.3	Translation of <i>egl-1</i> mRNA is blocked by <i>bc275</i> 3'UTR mutation	55
2.4	Fluorescent CIRSPR/Cas9 mediated tracing of distinct DNA sequences.....	58
2.4.1	Generation of an inactivated dCas9::GFP::GFP fusion construct.....	58
2.4.2	The dCas9::GFP::GFP fusion is embryonic lethal in <i>C. elegans</i>	59

2.4.3	Telomere targeting sgRNAs.....	60
2.5	<i>In silico</i> identification of chromatin modifiers with potential impact on <i>egl-1</i>	63
2.5.1	Hypothesis: chromatin modifiers open the <i>egl-1</i> locus prior to first wave of cell death	63
2.5.2	Identification of chromatin modifiers by the <i>C. elegans</i> gene ontology	64
2.5.3	Quality control of candidate genes	66
2.5.4	Expression data analysis of chromatin associated candidate genes	68
2.5.5	Overrepresentation analysis of candidate groups associated with GO term “programmed cell death”	72
2.5.6	Overrepresentation analysis of the highest ranked candidate group	76
3	Discussion.....	80
3.1	Post-transcriptional regulation of <i>egl-1</i> by microRNAs	81
3.1.1	Repression of <i>egl-1</i> by miR-35 and miR-58 family microRNAs is enhanced by cooperative effect	81
3.1.2	miR-35 family and miR-58 family microRNA activity equilibrates lineage-specific differences in <i>egl-1</i> transcription	83
3.1.3	The miR-35 and miR-58 families of microRNAs buffer <i>egl-1</i> expression during embryonic development	84
3.2	Analysis of the CRISPR/Cas9 generated <i>egl-1</i> 3'UTR variant <i>bc275</i>	86
3.2.1	Modification of the <i>egl-1</i> locus by CRISPR/Cas9	86
3.2.2	The gene variant <i>egl-1</i> (<i>bc275</i>) leads to increased mRNA stability and a general block in apoptosis	87
3.3	Fluorescent CRISPR/Cas9 mediated tracing of distinct DNA sequences	89
3.3.1	Future applications	90
3.4	<i>In silico</i> identification of chromatin modifiers with potential impact on <i>egl-1</i>	92

3.4.1	Chromatin factor identification and quality control	92
3.4.2	Future experiments	93
4	Material and Methods	95
4.1	Strains and maintenance	96
4.2	Cloning of the <i>egl-1</i> 3'UTR reporters with microRNA binding-site variants	98
4.3	Single copy integration of <i>egl-1</i> 3'UTR reporter	99
4.4	4D microscopy and lineage analysis of <i>C. elegans</i> embryonic development	102
4.5	Single-molecule RNA fluorescent in situ hybridization	102
4.5.1	Probes	102
4.5.2	Sample preparation	104
4.5.3	Imaging and quantification	104
4.6	CRISPR/Cas9 and sgRNAs	105
4.6.1	Cloning sgRNAs	105
4.6.2	Repair templates	106
4.6.3	Cloning <i>dCas9::gfp::gfp</i> fusion construct	106
4.6.4	Microinjection	109
4.7	Bioinformatic methods and tools	110
4.7.1	APIs and requested data	110
4.7.2	Expression data and analysis	110
4.7.3	Panther	112
5	References	113
6	List of figures and tables	142
7	Curriculum vitae	146

Informative abstract

During *C. elegans* development 131 somatic cells of the hermaphrodite die in a reproducible and invariant pattern. These deaths are triggered by the central apoptotic pathway, which is comprised of four components. EGL-1 a BH-3 only protein, which functions as the integrator of apoptotic stimuli. CED-9, a BCL-2-like protein, located at the outer mitochondrial membrane releases an asymmetric CED-4 dimer upon binding of EGL-1. The CED-4 dimer assembles into an octameric apoptosome which mediates the maturation of the zymogen proCED-3 into the catalytical active CED-3. The active CED-3 induces a downstream cascade which leads to cell death.

In the first part, I will present data which elucidate post-transcriptional regulation by microRNAs of the key integrator of the central apoptotic pathway, *egl-1*. I show that the miR-35 family and miR-58 family of microRNAs cooperate on repression of the *egl-1* mRNA by binding to conserved binding sites within the 3'UTR of *egl-1*. Loss of miR-35 family leads to abnormally large cell corpses in mothers or sisters of cells destined to die within the developing embryo. These cell death abnormalities are further enhanced by miR-58 family loss in double knockout mutants. Furthermore, the impact of the miR-35 and miR-58 family on *egl-1* expression is lineage dependent. This supports the overall proposal that lineage specific *egl-1* translational activation only triggers the central apoptotic pathway in daughter cells programmed to die.

Second I will present an insertion mutation within the *egl-1* 3'UTR which was generated by co-CRISPR/Cas9 directed insertion mutagenesis. The insertion within the *egl-1* 3'UTR leads to a change in the predicted 3'UTR folding and weakening of the inherent predicted stability. This is reflected in a complete absence of cell deaths throughout the entire embryonic development in a lineage independent manner. Further, in later stages of embryonic cell death absence is reflected in a *ced*-like phenotype. Using single molecule RNA FISH, it becomes apparent that *egl-1* mRNA is present within the embryo but is not translated into a functional protein. Quantitative analysis of

egl-1 mRNA within the RID lineage shows that the level of *egl-1* mRNA in the mutant is sufficient to trigger cell death. This cell death inducing threshold is also already crossed within the RID neuroblast. This indicates that post-transcriptional regulation of *egl-1* mRNA via 3'UTR is at least partially blocked. Interestingly, the *egl-1* mRNA turnover within the RID neuron seems to be independent from the 3'UTR and maintains wildtype levels.

In the third part, I focus on the adaptation of the CRISPR based visualization of distinct genomic loci, shown in mammals, to *C. elegans*. I therefore generated a catalytically inactive mutant of the Cas9 gene fused to two codon-optimized GFPs. These GFP differ in the codon usage to avoid unwanted recombination. Unfortunately, the dCas9 gene has a high toxicity when expressed within *C. elegans* at a level which is defined by the promoter. Besides this I show that the dCas9::GFP::GFP fusion is located in the nucleus and accumulates within nucleoli, which has been shown before. I conclude that the dCas9::GFP::GFP fusion is in general functional but needs a more precise expression, both in duration and level, to avoid toxic side effects.

Finally, identification of genes which have an impact on apoptosis is a laborious and time intensive task, especially for apoptosis defects or abnormalities within embryonic development. Within an *in-silico* analysis I identify chromatin-modifying genes which have a potential impact on key integrator of apoptotic stimuli, *egl-1*. The analysis is based on the *C. elegans* gene ontology and expression data available for defined stages during embryonic development. For prediction of the potential impact, I compare normalized expression patterns of candidates to the expression of *egl-1* which starts peaking close to the first wave of cell death. Gene ontology data is extracted via an API of the wormmine database and is subsequently processed within a self-developed Python application.

1 Introduction

1.1 *Caenorhabditis elegans* – a model organism

The nematode *C. elegans* is a small round worm, with about ~1100 µm in length and ~65 µm in diameter in the adult stage (Figure 1A). *C. elegans* can be found in nutrient- and bacteria-rich environments like rotten plants, yet the natural habitat is unknown (Kiontke and Sudhaus, 2006). *C. elegans* is vermiform, unsegmented and bilaterally symmetrical, furthermore it is transparent throughout the entire life cycle including the eggshell.

The lifecycle is trisected into embryogenesis, larval stages L1, L2, L3 and L4 (including Dauer stage) and the adult stage, it is traversed in 3 days at 25 °C. The egg of the developing worm has a dimension of ~50 µm in length and ~30µm in diameter (Figure 1B).

Embryonic development takes about 840 minutes at 22 °C, from first cell division until hatching of the worm. Further the development is subdivided into a proliferation and a morphogenesis/organogenesis stage (von Ehrenstein and Schierenberg, 1980; Sulston *et al.*, 1983; Wood, 1988). The proliferation stage is split into *in utero*, comprised of zygote formation until generation of founder cells (~30 cells; ~150 min) and *ex utero* development, including the majority of cell divisions and gastrulation (~558 cells; ~350 min) (Bucher and Seydoux, 1994). Organogenesis and morphogenesis largely overlap and are mainly described by the morphological changes the embryo undergoes (~350 min to ~840 min). During the morphogenesis/organogenesis stage the embryo elongates, triples in length and tissues and organs are fully generated (von Ehrenstein and Schierenberg, 1980; Sulston *et al.*, 1983). *C. elegans* complete cell lineage has been determined and is invariant between animals (Sulston, 1976; Sulston and Horvitz, 1977; Sulston *et al.*, 1983), making it possible to track single cells throughout the development.

C. elegans has two sexual forms (Figure 1A), the male nematode produces sperm and occurs with a frequency of 0.2 % naturally. The egg-laying and sperm-producing hermaphrodite can have up to 300 eggs of self-progeny. Both sexes have 5 diploid sets of autosomes arranged as linkage groups (LG) LG I to LG V. Genetically they differ in the sex chromosome LG X, one in the male and two in the hermaphrodite. The genome of *C. elegans* is fully sequenced and is comprised of 97 megabases (Mb) encoding for ~21,000 genes of which 38 % have a predicted ortholog in humans (Equence *et al.*, 1998; Shaye and Greenwald, 2011).

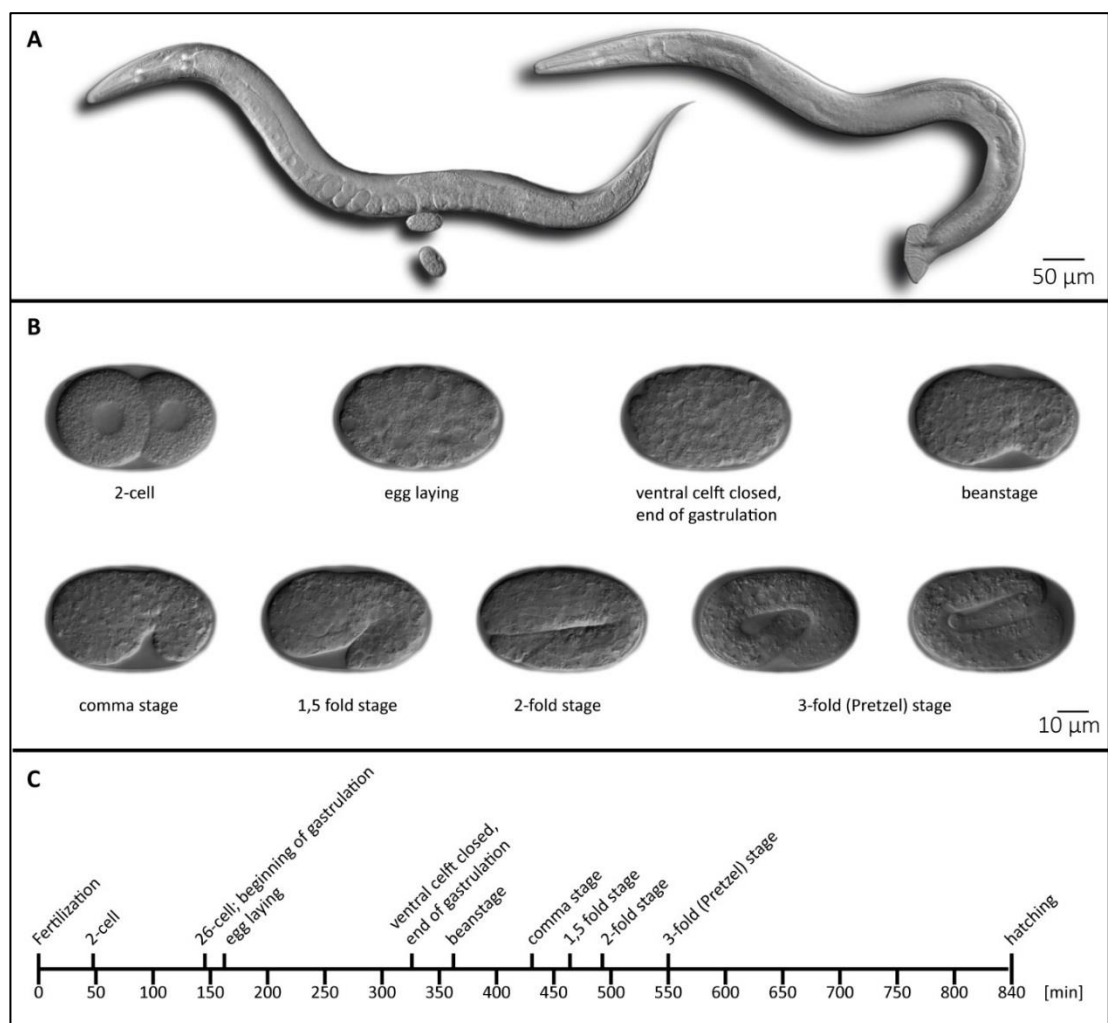


Figure 1. Representative images of *C. elegans*.

A) DIC images of adult *C. elegans* wild-type hermaphrodite (left) and male (right). (images by Nadin Memar and myself) **B)** DIC images *C. elegans* embryos at characteristic stages during

development (own images). **C)** Embryonic development of *C. elegans* from fertilization (0 minutes) to hatching (~ 840 minutes), time [min] shown on horizontal axis. Adapted from wormatlas.org (Wormatlas.org, 2017) based on (von Ehrenstein and Schierenberg, 1980; Sulston et al., 1983; Wood, 1988; Bucher and Seydoux, 1994; Chin-Sang and Chisholm, 2000).

1.2 Programmed cell death

Programmed cell death (PCD) has been studied extensively over the years and can be found throughout all domains of life and kingdoms. It was first observed in 1842 by Carl Vogt who saw the removal of notochord cells while studying the metamorphosis of the common midwife toad (*Alytes obstetricians*) (Vogt, 1842). In the past programmed cell death was defined by distinct changes of morphologies of dying cells and subdivided in apoptosis, necrosis and mitotic catastrophe (Kroemer *et al.*, 2009). Today a more functional classification based on molecular markers defines cell death in subcategories. In 2012 The Nomenclature Committee on Cell Death proposed the definitions of extrinsic apoptosis and intrinsic apoptosis which can be caspase dependent or independent, regulated necrosis, autophagic cell death, mitotic catastrophe and a tentative definition of other cell death (anoikis, cornification, entosis, netosis, parthanatos, pyroptosis) (Galluzzi *et al.*, 2012).

1.2.1 Apoptosis

Apoptosis refers to removal of cells by regulated killing which can be induced by intrinsic or extrinsic stimuli. Removal of structures which are no longer necessary is one key role of apoptosis. Further, structures only required by males or females (sexual dimorphism) are also eliminated by apoptosis; in the mammalian system the Müllerian ducts and Wolffian ducts are both developed along with the gonads. Upon sex determination, the Müllerian ducts were degraded in the male while in the female the Wolffian ducts were removed. Another important role of apoptosis is tissue remodeling

or sculpting, for example during embryonic development the interdigital webs between fingers were removed via apoptotic processes (Lindsten *et al.*, 2000). During development of the mammalian central nervous system (CNS), overproduced cells are eliminated by apoptosis to adjust matching numbers of different cells in the CNS (Oppenheim, 1991; Barres and Raff, 1999). Harmful or critically damaged cells, for example DNA double strand breaks induced by UV radiation, were also killed by the apoptosis machinery (Martin and Cotter, 1991; Nelson and Kastan, 1994; Rieber and Rieber, 1994).

Altered apoptosis levels, both high and low, are therefore harmful and can lead to abnormal development. High levels of apoptosis can further lead to neurodegenerative diseases like Huntington (Petersén A *et al.*, 2001) (reviewed in: Martin, 2010), while low levels can cause autoimmune diseases like lupus (reviewed in: Fenton, 2015) and are associated with cancer development (Kerr, Wyllie and Currie, 1972; Minn *et al.*, 1995; Raffo *et al.*, 1995; Fulda, Meyer and Debatin, 2002; Benard *et al.*, 2014) (reviewed in: Conradt, 2009; Fuchs and Steller, 2011; Labi and Erlacher, 2015). Cells undergoing apoptosis show characteristic morphological changes, comprised of cell rounding, blebbing, cell shrinkage, nuclear fragmentation, chromatin condensation, DNA fragmentation and mRNA decay (Kerr, Wyllie and Currie, 1972). This sequence of events prevents damage to the surrounding tissue caused by harmful cell components and is highly regulated (reviewed in: Fuchs and Steller, 2011; Labi and Erlacher, 2015). For comparison, necrosis, a traumatic cell death, is always noxious to the surrounding tissue by releasing cellular components uncontrolled into the extracellular space (reviewed in: Elmore, 2007).

1.2.2 Apoptosis in *Caenorhabditis elegans* development

During development of the *C. elegans* hermaphrodite 131 out of 1090 cells are eliminated by apoptosis; these cells undergo programmed cell death (PCD) in an invariant temporal and spatial manner. While 113 of these cells die during embryonic

stage of development, the remaining 18 die post-embryonically. Approximately 220 minutes after the first cell division (at 22 °C) the first cell deaths occur, and this wave of deaths continues until ~620 minutes. These deaths precede an asymmetric cell division, resulting in a daughter cell destined to die and a surviving sister cell (Sulston and Horvitz, 1977; Sulston *et al.*, 1983). This specification phase of death is determined by upstream pathways which lead to the activation of the cell death program within the activation phase. Under the microscope the dying daughter is characterized by a smaller cytoplasmatic volume compared to the sister cell. In the activation phase the cell death program is started, resulting in the execution phase where the cell is disassembled and finally killed. Using differential interference contrast (DIC) microscopy the dead cell can easily be identified through its roundness and change in refractility, leading to a “button” like shape. The dead cell is engulfed by the neighboring cells and gets fully degraded (Figure 2). Notably, it has been recently demonstrated that the cell death program is already executed within the mother cell, at least in two cell types, and plays a critical role in asymmetric cell division. (Chakraborty *et al.*, 2015; Lambie and Conradt, 2016; Mishra, Wei and Conradt, 2018). It should be noted that aspects of the specification, activation and execution phase may occur in parallel and not strictly sequentially (Hoeppner, Hengartner and Schnabel, 2001; Reddien, Cameron and Horvitz, 2001; Chakraborty *et al.*, 2015; Conradt, Wu and Xue, 2016; Lambie and Conradt, 2016).

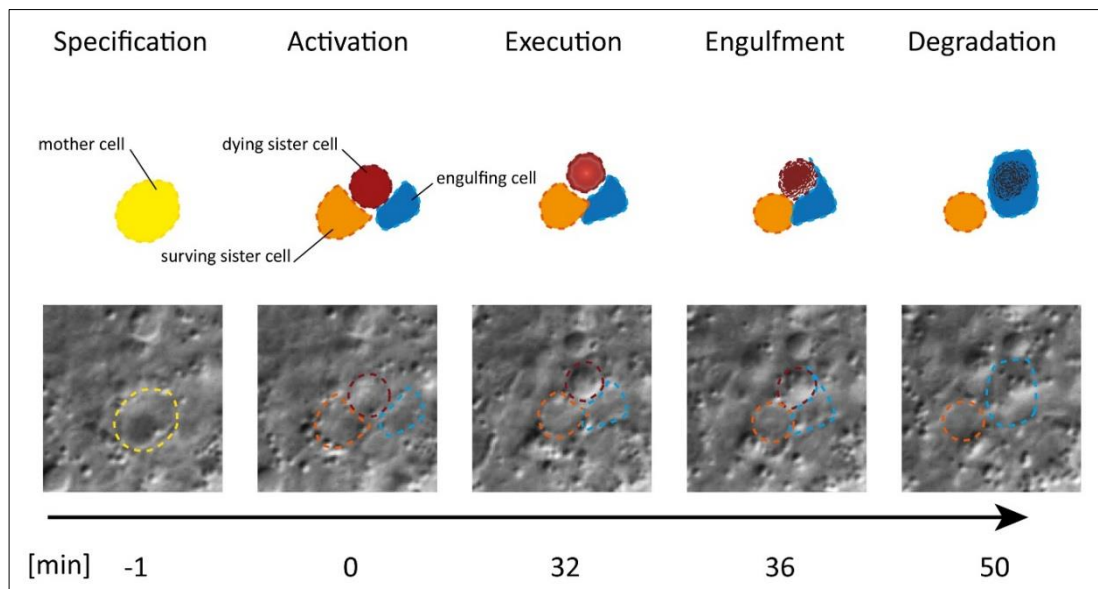


Figure 2. Phases of apoptosis in *C. elegans*.

In *C. elegans* a mother cell (yellow) divides into one daughter undergoing programmed cell death (red) and one daughter which is going to survive (orange). While the surviving daughter persists in size and shape, the dying daughter has inherited a smaller cytoplasmatic volume (0 min) and after 32 minutes rounds up and forms "button"-like shape visible under the DIC microscope. After 50 min, the cell gets engulfed by surrounding neighbors (own images and schematics).

1.2.3 The central cell death pathway in *Caenorhabditis elegans*

Genetically these 131 developmental cell deaths are induced by the central cell death pathway, which is comprised of four members. EGL-1 (egg laying defective) and CED-9, CED-4 and CED-3 (cell death abnormal). Loss of function (lf) mutations in either EGL-1 (Conradt and Horvitz, 1998), CED-3 or CED-4 lead to survival of nearly all 131 cells, which normally die during embryonic development; these genes therefore have a pro-apoptotic function (Ellis and Horvitz, 1986). In CED-9 gain-of-function (gf) mutants, apoptosis is blocked and contrary CED-9(lf) mutants show excessive cell death, resulting in an anti-apoptotic function of CED-9 (Hengartner *et al.*, 1992). Epistasis studies have revealed that the pathway is triggered by the most upstream factor, EGL-1

(Conradt and Horvitz, 1998). The direct pathway of apoptosis induction, therefore, starts with the presence of EGL-1. EGL-1 binds to the surface of the downstream factor CED-9, a BCL-2-like protein, located at the outer mitochondrial membrane. (Hengartner *et al.*, 1992; Hengartner and Horvitz, 1994). Notably, CED-9 shares significant sequence homology with Bcl-2 and Bcl-xL, mammalian antiapoptotic proteins (Hengartner and Horvitz, 1994). Negatively regulated by CED-9 an asymmetric dimer of CED-4 (APAF-1-like caspase activator) is bound to a specific binding site (Hengartner and Horvitz, 1994; Chinnaiyan *et al.*, 1997; James *et al.*, 1997; Spector *et al.*, 1997; Wu *et al.*, 1997; Chen, 2000). The EGL-1 binding site on CED-9 is at a different location than the CED-4 dimer binding site. Binding of EGL-1 induces a conformational change within CED-9, which alters the CED-4 binding site of CED-9, leading to a release of the CED-4 dimer (Conradt and Horvitz, 1998; del Peso, González and Núñez, 1998; Chen, 2000; del Peso *et al.*, 2000; Yan *et al.*, 2004). The CED-4 dimers relocate to perinuclear membranes; there four CED-4 dimers assemble one octameric apoptosome (Wu *et al.*, 1997; Chen *et al.*, 2000; Qi *et al.*, 2010). The assembled apoptosome mediates the activation of the caspase CED-3, which is expressed as inactive zymogen (proCED-3) (Yuan and Horvitz, 1992; Chinnaiyan *et al.*, 1997; Irmeler *et al.*, 1997; Seshagiri and Miller, 1997; Wu *et al.*, 1997; Yang, 1998). The active CED-3 in turn then induces a downstream cascade resulting in cell death. (Ellis and Horvitz, 1986; Greenwald and Horvitz, 1980) (Figure 3). For example, the exposure of phosphatidylserine on the cell surface is initiated by CED-3 mediated activation of CED-8 (Xkr8-like protein) (Stanfield and Horvitz, 2000; Suzuki *et al.*, 2013; Y.-Z. Chen *et al.*, 2013).

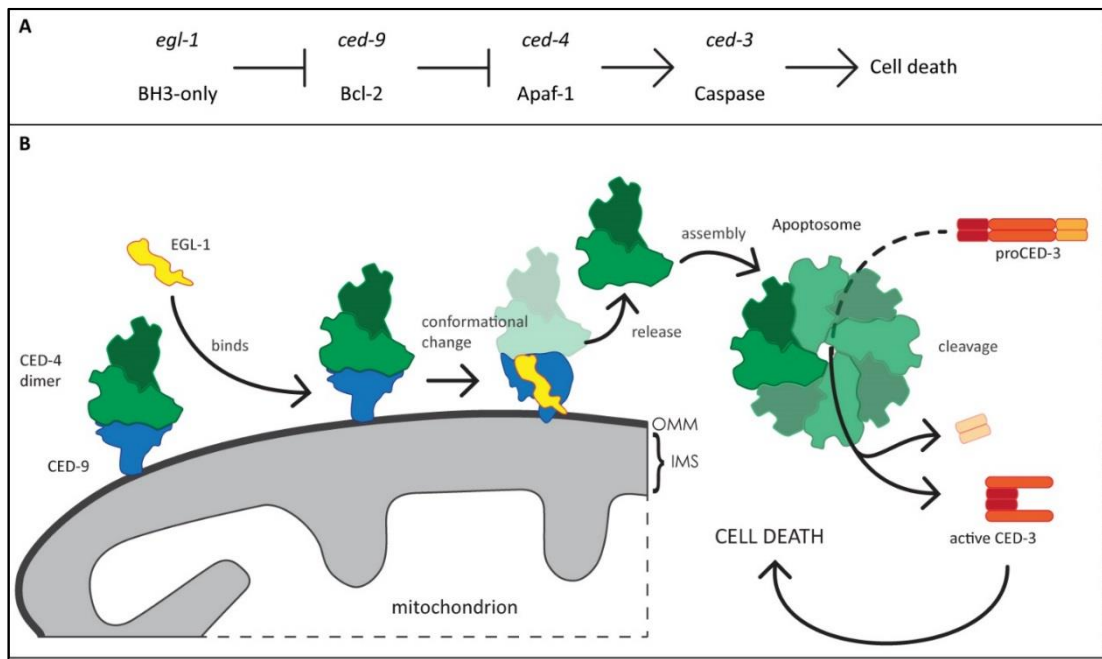


Figure 3. Representation of the central cell death pathway in *C. elegans*.

A) Members of the central cell death pathway in *C. elegans* (top) and their human homologs (bottom). **B)** CED-4 is bound as asymmetric dimer to CED-9 which anchors at the outer mitochondrial membrane (OMM). Upon binding of EGL-1 to CED-9, CED-9 undergoes a conformational change releasing the CED-4 dimer. CED-4 then assembles to an octameric apoptosome, which cleaves proCED-3 resulting in active CED-3 which leads to cell death. Own schematics based on Woo *et al.*, 2003; Yan *et al.*, 2004; Conradt and Xue, 2005; Huang *et al.*, 2013; references for intermediate steps see detailed explanation in 1.2.3)

1.2.4 EGL-1 key integrator of apoptotic stimuli

During *C. elegans* development exactly 131 cells which are programmed to die have a high level of *egl-1* activity, while in the surviving 959 cells *egl-1* activity is low or absent (Conradt and Horvitz, 1999; Thellmann, Hatzold and Conradt, 2003; Liu *et al.*, 2006; Hatzold and Conradt, 2008; Potts, Wang and Cameron, 2009; Hirose, Galvin and Horvitz, 2010; Winn *et al.*, 2011; Hirose and Horvitz, 2013; Wang *et al.*, 2015). Originally, *egl* genes were identified by screens in *C. elegans* for egg-laying (*egl*) defective mutants (Trent, Tsuing and Horvitz, 1983; Ellis and Horvitz, 1986). In *egl-1* gain of function (*gf*) animals this leads to the inappropriate death of the hermaphrodite-specific neurons (HSNs), which are required for egg laying (Conradt and Horvitz, 1999).

Genetically *egl-1* encodes for a 91 amino acid long protein, containing a nine amino acid long Bcl-2 homology region 3 (BH3), but not the BH1, BH2, or BH4 domain (BH3-only protein). Through the BH3 domain EGL-1 is capable of interacting with the *C. elegans* core BCL-2 family member, CED-9 (Conradt and Horvitz, 1998) (see 1.2.3). Within the genome and close to the transcription unit of *egl-1*, *cis*-regulatory elements are located, which are conserved among Caenorhabditis species (Figure 4). Mutations in these *cis*-regulatory elements lead to changes in *egl-1* expression and can alter the cell death penetrance and pattern (Conradt and Horvitz, 1999; Hirose and Horvitz, 2013). *Trans*-acting factors which directly influence *egl-1* expression (and therefore cell death) by direct interaction with *cis-elements* have been identified (Conradt and Horvitz, 1999; Thellmann, Hatzold and Conradt, 2003; Liu *et al.*, 2006; Potts, Wang and Cameron, 2009; Hirose, Galvin and Horvitz, 2010; Winn *et al.*, 2011; Hirose and Horvitz, 2013; Wang *et al.*, 2015) (Figure 4).

Going back to the HSNs, it was shown that TRA-1 a GLI-like zinc-finger transcription factor directly interacts with a cis-regulatory element ~5,6 kb downstream of the *egl-1* transcription unit. In *C. elegans* TRA-1 promotes female development and acts as terminal regulator of somatic sexual fate (Zarkower and Hodgkin, 1993; Zarkower, 2006). The *trans cis* interaction leads to repression of *egl-1* transcription and ultimately to cell survival. Since, TRA-1 activity is high in hermaphrodites, this leads to HSN survival. The opposite is true for males; here TRA-1 activity is low and ultimately leads to cell death of the HSNs (Conradt and Horvitz, 1999). This process is an interesting example of sexually dimorphic cell death specification. Furthermore, it has been shown that a co-repressor complex composed of PLZF (promyelocytic leukemia zinc finger protein)-like zinc-finger transcription factor, TRA-4, the histone deacetylase, HDA-1, and the histone chaperone, NASP-1, as well as the DRM (DRM, DP, Rb and MuvB) complex are required for efficient TRA-1-dependent repression of *egl-1* transcription in the HSNs in hermaphrodites (Grote and Conradt, 2006; Harrison *et al.*, 2006; Large and Mathies, 2007). These complexes are associated with chromatin remodeling, which shows that epigenetic gene regulation is important for *egl-1* transcriptional control.

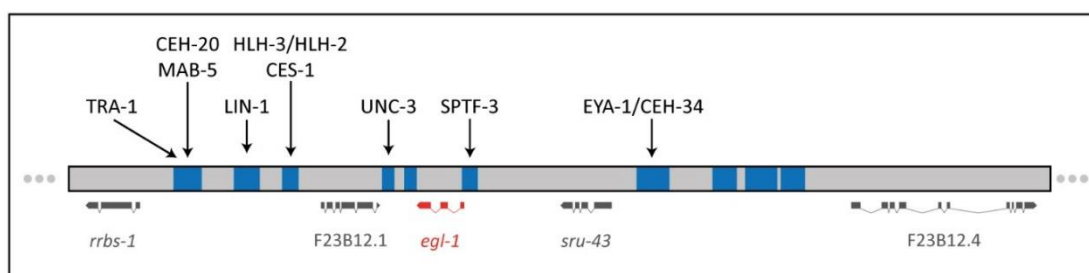


Figure 4. Regulation of *egl-1* transcription.

Depicted as a grey bar is part of chromosome V showing the *egl-1* as well as upstream and downstream transcription units. Blue regions represent conserved sites among *Caenorhabditis* species. *cis*-elements and their known corresponding transcription factors are shown in black, arrows indicating their binding site. Image adapted from Conradt, Wu and Xue, 2016

1.3 Chromatin structure, remodeling influence gene expression

1.3.1 Chromatin structure

Chromatin is the hierarchical structure of DNA, noncoding RNAs complexed with proteins. Its primary functions are the compaction of DNA, prevention of DNA damage, ordering the DNA macromolecule to allow mitosis and controlling gene expression. On the first level, the chromatin organization is comprised of ~146-147 base pairs (bp) of DNA, noncoding RNAs and associated proteins (Finch *et al.*, 1977; Carter and Jr., 1978; Luger *et al.*, 1997; Richmond and Davey, 2003). These are wrapped around histone octamer forming a nucleosome. The octameric structure of histones consists of two heterodimers of H2A and H2B together with H3 and H4 heterodimers (Finch *et al.*, 1977; Carter and Jr., 1978; Luger *et al.*, 1997). All histone proteins share 3 alpha helices (A1 to A3) and 2 disordered linkers (L1 and L2), further they are rich in the lysine and arginine amino acids to facilitate binding to the negatively charged DNA backbone (Arents and Moudrianakis, 1995; Luger *et al.*, 1997). Nucleosomes are connected through linker DNA of 10 – 90 bp in length, this connection can be observed under the electron microscope as “beads on a string” structure (Olins and Olins, 1974; Li, 1975; Oudet, Gross-Bellard and Chambon, 1975). Together with histone H1, the “beads on a string” coil into a helical secondary structure of 30nm diameter (30nm fiber) (Li, 1975; Wigler and Axel, 1976; Woodcock and Dimitrov, 2001). Long-distance contacts between secondary structure fibers form the tertiary form of chromatin structure, which can be found as specialized chromatin loops and chromosomes in metaphase (Woodcock and Dimitrov, 2001; Woodcock and Ghosh, 2010). Based on the structure chromatin is described in two varieties as euchromatin and heterochromatin. Historically, these two forms were distinguished by the intensity of staining. In interphase nuclei, the loosely packed euchromatin appears light-colored compared to the densely-packed heterochromatin (Heitz, 1928; Passarge, 1979). Euchromatin is active in transcription of genes, these genes are arranged around nucleosomes (11 nm fiber / secondary

structure of chromatin) and have a high GC-content compared to genes in the heterochromatin (Bickmore and Sumner, 1989). Heterochromatin contains genes which are mostly transcriptionally silenced; it is normally densely packed around evenly distributed nucleosomes forming a higher order structure (30 nm fiber / tertiary structure) (Tremethick, 2007). Further, heterochromatin is subdivided into a facultative and constitutive state. Facultative chromatin can cycle between euchromatin and heterochromatin by rearrangement of the underlying chromatin structure (Trojer and Reinberg, 2007). Constitutive heterochromatin remains transcriptionally silent and consists mainly of tandem repeats maintaining the structural integrity of chromatin (Le *et al.*, 2004; Craig, 2005).

1.3.2 Telomeres protect the end of linear chromosomes

The end of linear chromosomes is protected by telomeres, comprised of repetitive DNA repeats built of a species-specific conserved motif. In vertebrates the length of the repeating motif ranges from 9 – 15 kb, sharing the consensus sequence (5'-TTAGGG-3') (Meyne, Ratliff and Moyzis, 1989). In *C. elegans* however, chromosomes terminate after 4 – 9 kb of tandem repeats (5'-TTAGGC-3') (Wicky *et al.*, 1996). These repeats are associated with the sheltering protein complex forming a telomere stabilizing T-loop (Griffith *et al.*, 1999; Stansel, de Lange and Griffith, 2001). Preventing the telomere from being recognized as a DNA double strand break, this further resolves the end replication problem as well as preventing chromosomal fusion (McClintock, 1938; Muller, 1938; Nikitina and Woodcock, 2004).

1.3.3 Chromatin modifications

The chromatin architecture allows dynamic modifications to access transcriptionally inactive genes (activation) or inactivate transcriptionally active genes (silencing). This chromatin remodeling is achieved by histone modification and ATP-dependent chromatin remodeling complexes. Modification of histones comprises acetylation, methylation, phosphorylation, ubiquitination, ribosylation as well as sumoylation, established by a variety of chromatin modifying complexes. ATP-dependent chromatin remodeling complexes recognize specific histone modifications and restructure nucleosome positions and density. In *C. elegans* histones share, compared with humans, 80 % amino acid sequence identity, and H3 and H4 have been shown to be nearly identical (> 95%) to human H3 and H4 proteins (Vanfleteren, van Bun and Van Beeumen, 1987; Vanfleteren, Van Bun and Van Beeumen, 1987; Vanfleteren, Van Bunt and Van Beeumen, 1989). It is predicated that *C. elegans* has homologs of all mammalian histone modifying enzymes (Cui and Han, 2007). Furthermore, *C. elegans* has many homologs to chromatin remodeling complexes present in mammals, including SWI/SNF (Sawa, Kouike and Okano, 2000), NURD/CHD (Lu and Horvitz, 1998; Solari and Ahringer, 2000), ISWI (Andersen, Lu and Horvitz, 2006) and SWR1 (Ceol and Horvitz, 2004). Several of these chromatin modifying factors play an essential role in cellular and developmental processes of *C. elegans* (Cui and Han, 2007).

1.4 Post-transcriptional regulation of gene expression

1.4.1 Mechanisms of post-transcriptional regulation

While chromatin structure and modifications regulate gene transcription, post-transcriptional regulation is defined as the control of gene expression at the RNA level. As the term post indicates, this type of regulation takes place between transcription and translation of a gene. Focusing on mRNA, post-transcriptional regulation comprises several processes which lead in the end to translation or decay of mRNA. The mRNA is normally comprised of five elements, at the 5' end the mRNA is protected by a cap structure. Subsequently followed by 5'UTR also referred to as leader sequence. Attached to the 5'UTR and usually starting with a start codon is the coding sequence (CDS), which contains the part of the gene which encodes for the final protein and is normally terminated by a stop codon. The 3'UTR spans from the stop codon to the beginning of the polyadenylated tail of the mRNA. All elements of the mRNA can be modified by proteins and complexes, summarized as post-transcription regulation. However, in eukaryotes three processing steps are mainly responsible for the maturation of pre-mRNA into mRNA.

Firstly the 5' cap structure, which gets co-transcriptionally attached to nascent mRNA by applying a 7-methylguanosine cap at the 5'-end of the transcript linked to the first nucleotide (Shatkin and Manley, 2000; Moteki and Price, 2002). The main functions of the 5' cap are its protective function against 5' to 3' exonuclease cleavage (Furuichi, LaFiandra and Shatkin, 1977; Shimotohno *et al.*, 1977), signaling for protein factors essential for pre-mRNA splicing, regulating the nuclear export (Visa *et al.*, 1996; Lewis and Izauride, 1997) and promoting cap-dependent translation (Dever, 1999; Merrick, 2004). Hydrolysis of the 5' cap (mRNA decapping) promotes quick mRNA decay (Meyer, Temme and Wahle, 2004), which is mediated by DCP1, DCP2 and DcpS in *C. elegans* (Lall, Piano and Davis, 2005) and is one key elements of regulation.

Secondly, removal of introns and noncoding regions of the transcribed pre-mRNA is catalyzed by the spliceosome, a complex of small nuclear ribonucleoproteins (Will and Lührmann, 2011). It is estimated that 95 % of all multi-exon genes are further alternatively spliced leading to a varying exon composition of the same mRNA and ultimately to different proteins, which influences gene translation in a quantitative manner (Kapranov *et al.*, 2002; Harrow *et al.*, 2006). Changes in the 5' or 3'UTR lead to changes in mRNA stability, localization and differential translational control (Stamm *et al.*, 2005; Hughes, 2006). Further, alternative splicing can also induce premature termination codons (PTC) leading to nonsense-mediated mRNA decay (NMD) (Lewis, Green and Brenner, 2003; McGlincy and Smith, 2008). Interestingly, splice variants of the human Bcl-x gene have two opposite activities, while Bcl-x(L) is antiapoptotic, Bcl-x(s) has a proapoptotic function (Boise *et al.*, 1993).

Polyadenylation of nascent mRNA is another element of post-transcriptional regulation. It is established by the processive polyadenylation complex in the nucleus of eukaryotes (Bienroth, Keller and Wahle, 1993). The 3' end of the mRNA, which often contains a polyadenylation signal AAUAAA gets cleaved, and the poly(A) tail gets attached by the polyadenylate polymerase. The poly(A) tail is approximately 200 to 250 nucleotides long, protecting the mRNA from enzymatic degradation in the cytoplasm (Wickens, 1990; Wahle and Keller, 1992; Decker and Parker, 1994). The export of the mRNA from the nucleus and translation as well as transcription termination is also supported by the poly(A) tail (Lewis, Gunderson and Mattaj, 1995; Coller, Gray and Wickens, 1998).

These major modifications to pre-mRNA and their associated interaction partners contribute to mRNA stability leading in the end to translation or decay of a mRNA. On a broader scale, mRNAs with a short half-life can respond fast to changes in gene-expression compared to long living mRNAs (Ross, 1995). Differential mRNA turnover rates can contribute to changes in gene expression due to rapidly changing

developmental and/or environmental conditions. On the other hand, they can maintain levels of translatable mRNA rigidly. This opens a window for a prolonged translation of highly expressed genes, but also for translational delayed genes, which can be stably stored (Guhaniyogi and Brewer, 2001).

1.4.2 Post-transcriptional regulation via the 3' untranslated region

Besides the three major regulators of post-transcriptional gene regulation, 5' capping, splicing and polyadenylation, mRNAs are also regulated via regulatory regions within the 3' untranslated region (3'UTR). The 3'UTR is located downstream of the coding sequence and is involved in mRNA localization, translation, polyadenylation, transcript stability and cleavage. 3'UTRs have regions which are highly conserved in the mammalian as well as in the *C. elegans* genome (Siepel *et al.*, 2005). The conservation is primarily based on one strand; this strengthens the role in post-transcriptional regulation (Xie *et al.*, 2005).

Many 3'UTRs contain microRNA response elements, sequences to which miRNAs bind and this will be explained in capture 1.5. Further, the stability of mRNA is influenced by *cis*-regulatory adenylate- and uridylylate (AU)-rich elements (ARE) (Jeffares, Penkett and Bähler, 2008). AREs are of 50 - 150 bp in length and contain multiple copies of the AUUUA motif (Chen and Shyu, 1995). ARE binding proteins (ARE-BPs) interacting with the AUUUA motif generally lead to degradation of the targeted mRNA (Chen and Shyu, 1995) but also stabilization is possible. For example, BCL-X_L expression is increased by stabilization of the mRNA through the ARE-BP nucleolin after UVA exposure. (Zhang, Tsapralis and Bowden, 2008). Most of the ARE-BPs are expressed in a tissue and cell type specific manner, with ARE secondary structure being important for their activity (Meisner *et al.*, 2004; Reznik and Lykke-Andersen, 2010). Another important aspect is the secondary structure of 3'UTR ; mutations altering the secondary structure are associated with various types of diseases for example β -thalassemia (Bilenoglu, Basak

and Russell, 2002) or type 2 diabetes (Xia, Bogardus and Prochazka, 1999) an overview is given by Chen et al. (Chen, Férec and Cooper, 2006).

1.5 Post-transcriptional regulation by MicroRNAs

1.5.1 MicroRNAs, biogenesis and function

MicroRNAs (miRNAs or miRs) are small non-coding RNA molecules found in plants, animals, fungi and viruses (Chen and Rajewsky, 2007). These often highly conserved miRNAs function in transcriptional and post-transcriptional regulation of gene expression (Bartel, 2009) by induction of messenger RNA (mRNA) degradation, inhibition of translation and induction of DNA methylation. It is estimated that the human genome may encode over 1000 miRNAs, which target about 60 % of all genes (Friedman *et al.*, 2009).

For targeting mRNAs in animals, partial complementarity is sufficient for miRNAs; these complementary regions are called “seed regions” and are about 6-8 nucleotides in length. This leads to the conclusion that miRNAs may have multiple different mRNA targets and vice versa. Complex combinatorial regulation of mRNAs is therefore a feature of miRNA dependent regulation of gene expression (Lewis, Burge and Bartel, 2005).

MicroRNAs were first characterized in 1993 by Victor Ambros *et al.* during studies of the gene *lin-14* in *C. elegans*. They found that a precursor from the *lin-4* gene matured to a 22-nucleotide RNA that is partially complementary to sequences in the 3'UTR of the *lin-14* mRNA and has the ability to block translation into protein (Lee, Feinbaum and Ambros, 1993). In *C. elegans* postembryonic development a *lin-14* gradient is established, controlling the transition from L1 to L2 larval stages, allowing the temporal control of gene expression to ensure correct timing of developmental events (Wightman *et al.*, 1991; Wightman, Ha and Ruvkun, 1993; Olsen and Ambros, 1999). Cancer development and prognosis are also associated with miRNAs, for example the *let-7* gene functions as a tumor suppressor by direct negative regulation of RAS/*let-60* (Johnson *et al.*, 2005) (reviewed in Lu *et al.*, 2005; Kusenda *et al.*, 2006).

The biogenesis of miRNAs can be divided into three steps comprising nuclear processing, nuclear export and cytoplasmic processing. During nuclear processing the miRNA gene is transcribed by RNA polymerase II and is capped at the 5' end while the 3' end is extended with a poly-adenosine tail (CAI, Hagedorn and Cullen, 2004; Lee *et al.*, 2004). This primary miRNA forms a hairpin structure and is recognized by DGCR8 (Pasha in invertebrates) which is associated with the RNA cutting enzyme Drosha (Lee *et al.*, 2003). This complex cuts the pri-miRNA leaving an overhang at the 3' end (Denli *et al.*, 2004; Gregory *et al.*, 2004; Han *et al.*, 2004). These pre-miRNA hairpins are exported into the cytosol via Exportin 5 mediated transport (Yi *et al.*, 2003; Bohnsack, Czaplinski and Gorlich, 2004; Lund *et al.*, 2004; Zeng and Cullen, 2004). The pre-miRNA hairpin is cleaved by the endonuclease Dicer (Bernstein *et al.*, 2001; Grishok *et al.*, 2001; Hutvagner *et al.*, 2001; MacRae *et al.*, 2006; Park *et al.*, 2011) and one miRNA strand is integrated into the RNA-induced silencing complex (RISC) known as the guiding strand (Tabara *et al.*, 1999; Hammond *et al.*, 2001; Okamura *et al.*, 2004; Song *et al.*, 2004; Kawamata and Tomari, 2010; Elkayam *et al.*, 2012). (Figure 5).

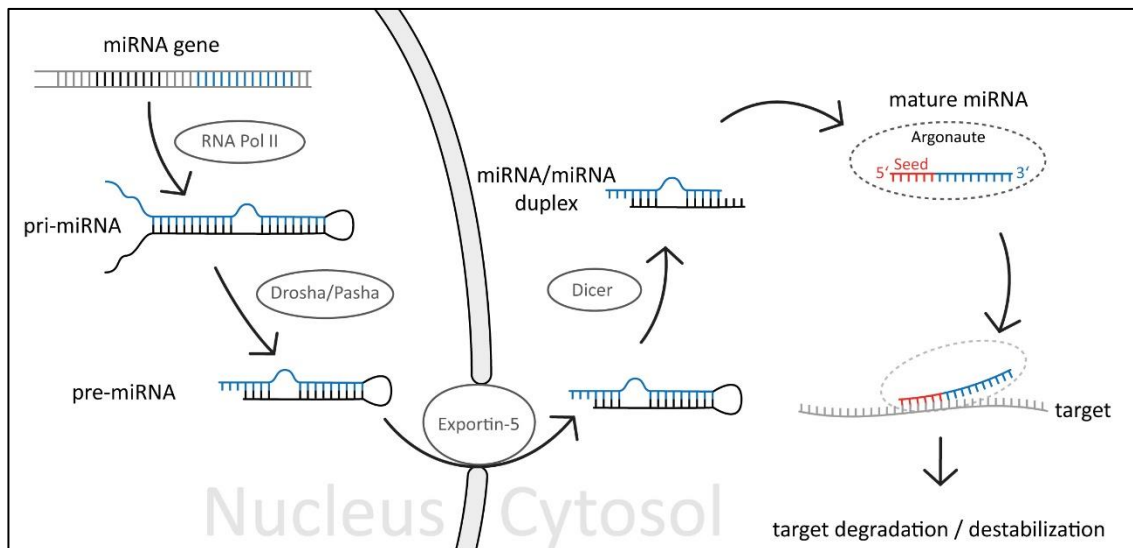


Figure 5. Biogenesis of microRNAs.

In the nucleus, the miRNA gene is transcribed by RNA polymerase II forming the primary miRNA. This is followed by the processing of the Drosha/Pasha complex into the precursor miRNA which is exported via Exportin-5 into the Cytosol. Here the Dicer complex processes the pre-miRNA into miRNA/miRNA duplex of which one strand is loaded onto an Argonaute family member (blue), while the other strand is degraded (black). The matured miRNA guides Argonaute to their target by base pairing with its seed sequence (orange), leading to target decay and/or destabilization. Own schematics based on Winter *et al.*, 2009; detailed references for intermediate steps in 1.5.1)

1.5.2 miRNA-mediated regulation of programmed cell death

The pro apoptotic BH3-only proteins Bim and PUMA have been shown to be targeted by various microRNAs repressing translation in many tissues. For example, the microRNA cluster miR-17~92 prevents maturation of Bim and synergizes with miR-106b~25 (Ventura *et al.*, 2008) enhancing the double knockout phenotype over the single knockout. An identical effect has been shown for PUMA; here miR-24 and miR-29 cause a synergistic effect in preventing PUMA induced cell death (Annis *et al.*, 2016).

With regards to cancer treatment, the role of miRNAs seems to be emerging. For example, depletion of miR-663 is sufficient to stop tumor progression in non-small lung cancer. Here miR-663 leads to mitochondrial depolarization and cell death by directly targeting PUMA (Fiori *et al.*, 2018). Interestingly, prevention of cell death is also used by the Epstein-Barr virus, insuring host cell survival upon infection by encoding the microRNA miR-BART5 preventing maturation of PUMA (Choy *et al.*, 2008). So far little is known about the post-transcriptional and translational control of EGL-1 in *C. elegans*. A previous study by Wu *et al.* showed that *egl-1* mRNA is targeted *in vitro* by the miR-35 and the miR-58 family miRNAs, potentially leading to translation repression by deadenylation (Wu *et al.*, 2010). Further, it has been shown that the miR-58 family is repressing targeted mRNAs efficiently and that loss of the whole miR-58 family leads to decreased germline apoptosis. While the study looked at an overall level of protein and mRNAs, a single target was not identified. (Subasic *et al.*, 2015).

1.5.3 miR-35-family and miR-58-family miRNAs in *C. elegans*

The *mir-35* family is comprised of eight members (*mir-35/36/37/38/39/40/41/42*). All members are similar in sequence and the microRNAs *mir-35* to *mir-41* are derived from the same genomic cluster. The *mir-42* is transcribed as part of the *mir-42-44* cluster, along with *mir-43* and *mir-44* which share no sequence similarity to the *mir-35* family. Both clusters are located on LGII lying 350kb apart. (Wormbase.org, 2017a, 2017b). The miR-35 family is highly expressed in oocytes and during early and mid-stage embryos (< 350 cells; ~8-9 round of cell division), with fast decreasing levels in later embryonic stages and nearly absent throughout L1 to L4 stages (Lau, 2001; Kato *et al.*, 2009; Stoeckius *et al.*, 2009; Alvarez-Saavedra and Horvitz, 2010; Wu *et al.*, 2010). In the adult stage the *mir-35* family is expressed with lower levels compared to embryonic stages (Lau, 2001; Kato *et al.*, 2009). Mutants that lost all members of *mir-35-41* (*nDf50*) and *mir-42* (*nDf49*) (referred as *mir-35* family mutants) arrest around the two-fold to three-fold stage of embryonic development and are not viable (Alvarez-Saavedra and

Horvitz, 2010). In contrast mutants which have only a deletion in *mir-35-41* (*nDf50*) are temperature sensitive for viability, with 50% survival at 15 °C and 1 % survival at 26 °C (Alvarez-Saavedra and Horvitz, 2010). They are further characterized by a reduced silencing of endogenous RNAi targets, a weakened hypoxic response, misregulation of RNA binding proteins and reduced fecundity (Massirer *et al.*, 2012; McJunkin and Ambros, 2014; Kagias and Pocock, 2015).

The *mir-58* family includes six members, *mir-58.1*(*mir-58*), *mir-80*, *mir-81*, *mir-82*, *mir-58.2* (*mir-1834*) and *mir-2209.1* which are homologous to the bantam family in *Drosophila*. *mir-58* is located on chromosome IV, *mir-80* on chromosome III and *mir-81* in close proximity to *mir-82* on chromosome X. The *mir-58* family is the most abundantly expressed miRNA family in *C. elegans*, covering > 30% of the total miRNAs (Jan *et al.*, 2011; Subasic *et al.*, 2015). miR-58.1, miR-80, miR-81 and miR-82 are highly abundant and are expressed with increasing levels from mid-stage-embryogenesis throughout the whole development. miR-58.2 and miR-2209.1 share a low abundance and are found with low levels at all stages of embryonic development (Stoeckius *et al.*, 2009; Alvarez-Saavedra and Horvitz, 2010; Isik, Korswagen and Berezikov, 2010; Wu *et al.*, 2010). Loss of a single miR-58 family member has no phenotypic effect but loss of the four highly abundant miR-58 family members [*mir-58.1*(*n4640*), *mir-80* (*nDF53*), *mir-81-82*(*nDF54*)] results in defects in locomotion and dauer formation and further reduce brood and body size (Alvarez-Saavedra and Horvitz, 2010).

1.6 CRISPR/Cas system

Studies on bacteria and archaea have identified an array of mechanisms to protect them against bacteriophages and viral invaders, ranging from inhibition of attachment over induced programmed cell death to cleavage of foreign DNA (Tock and Dryden, 2005; Barrangou *et al.*, 2007; Stern and Sorek, 2011; Jorgensen and Miao, 2015; Jorgensen, Rayamajhi and Miao, 2017). First observed in *Escherichia coli*, clustered regularly interspaced short palindromic repeats (CRISPR) and their CRISPR-associated (Cas) genes constitute a defense system providing adaptive immunity against invading genomic elements (Ishino *et al.*, 1987; Jansen *et al.*, 2002; Sorek, Kunin and Hugenholtz, 2008; Terns and Terns, 2011). These CRISPR/Cas systems induce a double strand break (DSB) leading to decay of targeted invading DNA.

The best studied Cas system is the type IIa system of *Streptococcus pyogenes* and its component CRISPR-associated protein 9 (*Sp* Cas9 or Cas9) (Anders *et al.*, 2014). Here, short stretches of foreign DNA are recognized by a complex of Cas1, Cas2, Csn2 and Cas9 which is associated with a *trans*-acting CRISPR (tracrRNA) (Datsenko *et al.*, 2012; Plagens *et al.*, 2012; Swarts *et al.*, 2012; Anders *et al.*, 2014; Nishimasu *et al.*, 2014; van der Oost *et al.*, 2014; Heler *et al.*, 2015; Wei, Terns and Terns, 2015). Upon selection of Cas9-mediated identification of a protospacer adjacent motif (PAM), the sequence close to their identification site is processed to a spacer-sized fragment (Stern *et al.*, 2010; Heler *et al.*, 2015). While still complexed with Cas proteins the spacer fragment is integrated into the CRISPR array by induced single strand breaks on each side of the repeats (Nuñez *et al.*, 2015). Transcription of these arrays results in pre-CRISPR-RNAs (crRNA) containing parts of the repeat sequence and a single spacer complementary to the foreign DNA (Deltcheva *et al.*, 2011). Along with the CRISPR array the tracrRNA encoded in close proximity is transcribed, forming tracrRNA:pre-crRNA duplexes bound by Cas9. Maturation of these duplexes involves RNase III in a first step while the second step remains elusive, resulting in the final mature crRNA

complexed with the tracrRNA and the endonuclease Cas9, forming the active silencing complex for targeting a specific foreign DNA sequence (Deltcheva *et al.*, 2011; Ratner, Sampson and Weiss, 2016).

1.6.1 Targeted genome modification

The utilization of Cas Ila system of *Streptococcus pyogenes* (*Sp* Cas9) and its programmable nucleases to generate targeted endogenous lesions had been used widely to edit genomes in many kingdoms (Datsenko *et al.*, 2012; Dickinson *et al.*, 2013; Mali, Esvelt and Church, 2013; Heler *et al.*, 2015; Paix *et al.*, 2015). In *C. elegans* and other organisms the CRISPR/Cas9 system for targeted genome modification consists of two main components: the cas9 gene and a concatenated chimeric form of crRNA and tracrRNA called single guide RNA (sgRNA) (Jinek *et al.*, 2012). While for the *Sp* Cas9 the PAM sequence 5'-NGG-3' is necessary, there are different PAM sites available which are broadening the range of possible targets (Karvelis, Gasiunas and Siksnys, 2017). The PAM site is therefore the limiting factor in designing sgRNAs since the target sequence must include the PAM sequence.

The CRISPR/Cas9 induced DSB can be repaired by either non-homologous end joining (NHEJ) or homology-directed repair (HDR). The NHEJ repair system is error-prone and induces with a high frequency insertions or deletions (indels) near the double strand break, causing mainly frameshift mutations or stop codons, which could lead to a knock-out of the gene of interest (Aravind and Koonin, 2001; Gu and Lieber, 2008; Lieber, 2010). The HDR can, by supplying a DNA homologous to the region next to the break site, be used to integrate designed sequences as well as, in close proximity, to replace or delete sequences (Bailis and Rothstein, 1990; Chen, Fenk and de Bono, 2013; Wu *et al.*, 2013).

1.6.2 Visualizing distinct DNA loci

While fluorescence in-situ hybridization allows sequence specific visualization of DNA and RNA in fixed samples, the CRISPR/Cas9 system was recently modified to allow DNA and RNA visualization in living cells (B. Chen *et al.*, 2013; Anton *et al.*, 2014; Nelles *et al.*, 2016). Therefore, point mutations in the Cas9 in both nuclease domains (RuvC and HNH) were introduced. The catalytically dead Cas9 (dCas9) is still capable of binding DNA via sgRNA targeting but lacks cleavage capability (Jinek *et al.*, 2012; Cong *et al.*, 2013). Fusing of the dCas9 to a fluorophore via a linker (including nuclear localization signal (NLS)), allows new perspectives on visualization of distinct loci in living cells and organisms. In mammalian cells, this approach has demonstrated its capability of visualizing loci, such as telomeres, containing many repeats of a single sgRNA (B. Chen *et al.*, 2013; Anton *et al.*, 2014) and also visualization of non-repetitive sequences by targeting through multiple sgRNAs (B. Chen *et al.*, 2013).

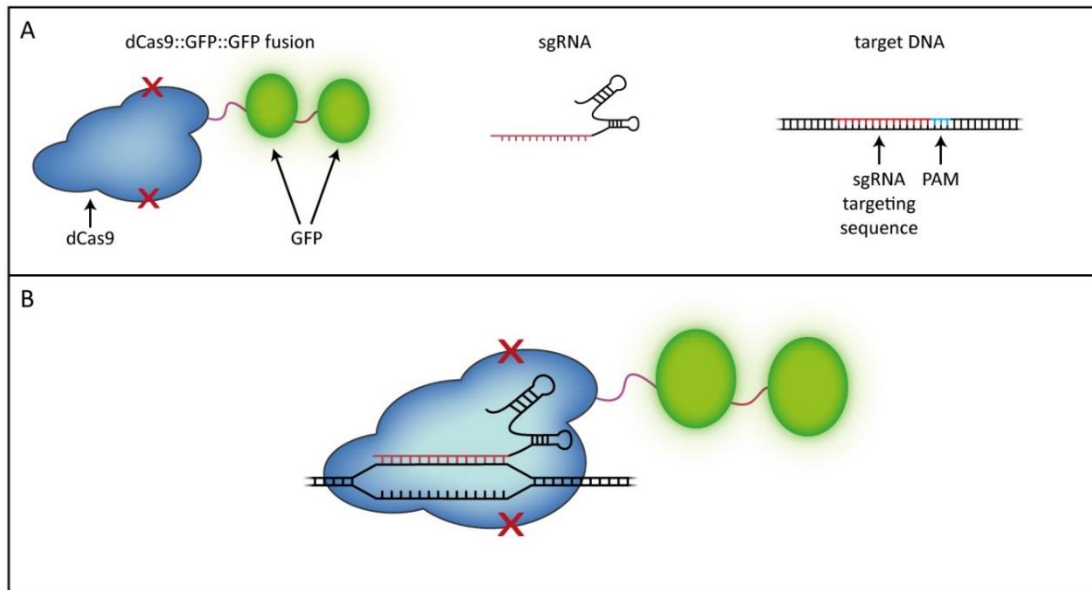


Figure 6. Labeling targeted DNA sequence with a dCas9::GFP::GFP/sgRNA complex.

A) Components of the targeted labeling; dCas9::GFP::GFP fusion protein (left); sgRNA (middle) comprised of the RNA scaffold (black) and the binding sequence (red); target DNA harboring the sgRNA targeting sequence (red) and the PAM site (blue). **B)** Schematic representation of the interacting components; dCas9::GFP::GFP fusion interacts with the sgRNA and is guided to the target sequence (own schematics based on Anton *et al.*, 2014).

1.7 Aim of this study

The aim of this study is to elucidate the regulation of key integrator of apoptotic stimuli, the BH-3 only protein, *egl-1*, during *C. elegans* embryonic development. While this is rather specific, the conservation of the central cell death pathway components in *C. elegans*, makes these results also of relevance for other organisms. This also includes the adaption of methods and development of tools *in vivo* and *in silico* which can be used for other research objectives.

Firstly, my colleagues Ryan Sherrard, Nadin Memar and I demonstrate that the miR-35 and miR-58 family of microRNAs target the *egl-1* mRNA *in vivo*, preventing unwanted cell death in mother and sister cells of a cell destined to die by directly repressing *egl-1*. Upon loss of the *mir-35* family genes a formation of large cell corpses in the developing embryos becomes apparent, which can be further enhanced by double knockout involving the *mir-58* family genes as well. The clear connection to central apoptotic pathway became obvious when knocking out *egl-1* or *ced-3* in *mir-35* family knockout background, which leads to a block in large cell corpse formation. Finally, by using fluorescent 3'UTR reporters it is shown that *egl-1* is directly targeted by the miR-35 family and miR-58 family at conserved sites mediating repression.

Then, in the second part I will examine two co-conversion CRISPR/Cas9 approaches. With the help of these approaches, I generated an *egl-1* 3'UTR insertion variant, which changes the predicted secondary structure. While there is no cell death visible throughout the entire *C. elegans* development, *egl-1* mRNA is present but not translated into a functional protein. The levels of the *egl-1* mRNA variant indicate that in a cell destined to die 3'UTR based repression of *egl-1* is dysfunctional.

In the third part I aimed to adapt the CRISPR based visualization of distinct genomic loci shown in mammals to *C. elegans*. I therefore generated a catalytically inactive mutant of the Cas9 gene fused to two codon-optimized GFPs which differ in their codon

usage. Unfortunately, the dCas9 gene induces a high toxicity when expressed within *C. elegans* at a high level. Nevertheless, I could show that the dCas9::GFP::GFP fusion is located in the nucleus and is potentially functional and I will discuss and recommend future optimizations.

Finally, identification of genes which have an impact on apoptosis is a laborious and time-intensive task, especially for apoptosis defects or abnormalities within embryonic development. Within an *in-silico* analysis I identify chromatin modifying genes ranked according to a hypothesized expression pattern of potential candidate genes. Therefore, I have developed a Python application which processes the *C. elegans* gene ontology and expression data which is available for embryonic development.

2 Results

2.1 Preface

The results presented in Section 2.2 were the combined effort of R. Sherrard, N. Memar, and me. This work was published in the January 15, 2017, issue of *Genes & Development*: Sherrard R, Luehr S, Holzkamp H, McJunkin K, Memar N & Conradt B (2017). miRNAs cooperate in apoptosis regulation during *C. elegans* development. *GenesDev* 31, 209–222.

For the results presented in section 2.3, hybridization of the smRNA FISH experiment was done by K. Ikegami. Counting of extra cells in the pharynx was done together with Barbara Conradt. The smRNA FISH probes for *egl-1* were designed by Ryan Sherrard.

The analysis presented in section 2.5 is based on available data from different data sources (see Materials and Methods). Since the tables generated in this analysis are not presentable on paper or pdf due to their size, they are available on the CD attached to this thesis.

2.2 Post-transcriptional regulation of *egl-1* by microRNA

2.2.1 The 3' untranslated region of *egl-1* contains conserved binding sites for microRNAs

The 3' untranslated region (3'UTR) of *egl-1* in *C. elegans* is comprised of 172bp. Comparing the 3'UTRs among *C. elegans*, *C. briggsae*, *C. remanei* and *C. brenneri* using TargetScanWorm (Jan *et al.*, 2011), shows highly conserved binding sites for the miR-35 family and miR-58 family (Figure 7 A) which were also found to associate with Argonaut ALG-1 in iPAR-CLIP experiments and are likely targets for the RNA-induced silencing complex (RISC) (Grosswendt *et al.*, 2014). The miR-35 binding site of *egl-1* is an exact match with the consensus seed sequence of the mature miR-35 family members (8-mer) in *C. elegans* and *C. brenneri*. While the miR-58 in *C. elegans* is an exact match with the seed sequence (8-mer); in other *Caenorhabditis* species the first and last base pair of the seed sequence is lacking. Interestingly, the miR-80 family is predicted as 6-mer seed in *C. elegans* lacking the first and last base pair, while the other *Caenorhabditis* species share a fully matching seed (8-mer).

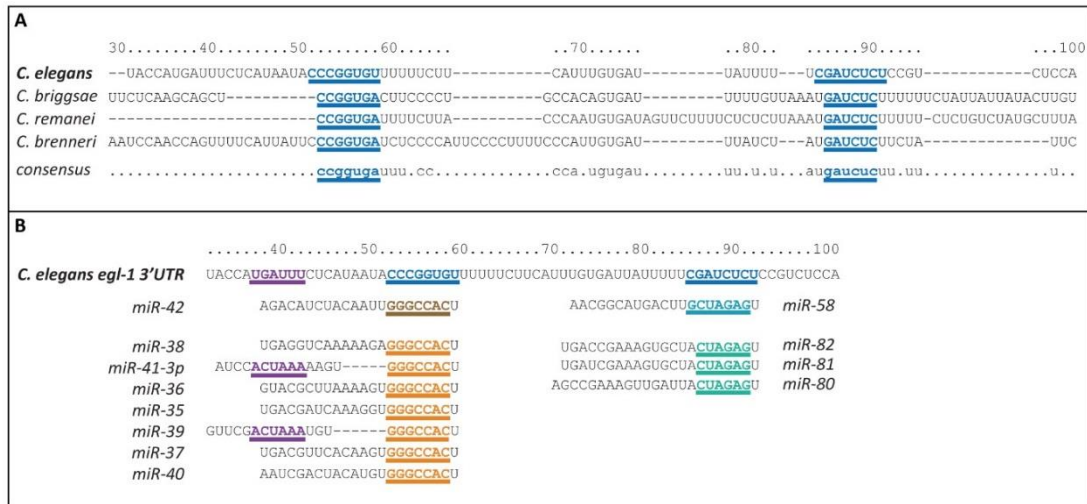


Figure 7. *egl-1* 3'UTR harbors conserved micro RNA binding sites

A) Conserved binding sites (blue) for *egl-1* 3'UTR of different *Caenorhabditis* species and their consensus sequence. **B)** Binding sites for miR-35 family in the *egl-1* 3'UTR in *C. elegans* (blue and violet) Seed sequences are shown for miR-42 (brown) miR-35/36/37/38/39/40 and miR-41-p3 (yellow). Additional consequential pairing is shown in purple. Seed sequences for miR-58 (light blue) and miR-80,81,82 (green). Data from TargetScanWorm (Jan *et al.*, 2011).

2.2.2 Abnormally large-cell corpses are present in *mir-35* family mutants

A phenotypic analysis using four-dimensional (4D) microscopy (Schnabel *et al.*, 1997, 2006) was performed to reveal cell death abnormalities in miR family mutants. I was focusing on the AB lineage, which is born with the first cell division and has a total of 98 cell deaths in hermaphrodites (95 in males) during embryonic development (Figure 8 B). The cell deaths occurring in the AB lineage were categorized in three waves, here a wave is defined as all cell deaths occurring within a time frame with temporal distance to cell deaths before and after the wave (Figure 8 B). In the AB lineage of wildtype (+/+) embryos at 20 °C the first wave of cell deaths begins at 250 min (close to the end of the 9th round of cell divisions) and ends at 260 min containing 13 cell deaths. The second wave is comprised of 55 cell deaths, starting within the 10th round of cell divisions at 360 min and ends prior to the beginning of the 11th round of cell division at 400 min. The third wave is comprised of 30 cell deaths occurring in the 11th round of cell division, starting from 425 min after the first cell division and ends at 430 mins (see Figure 8).

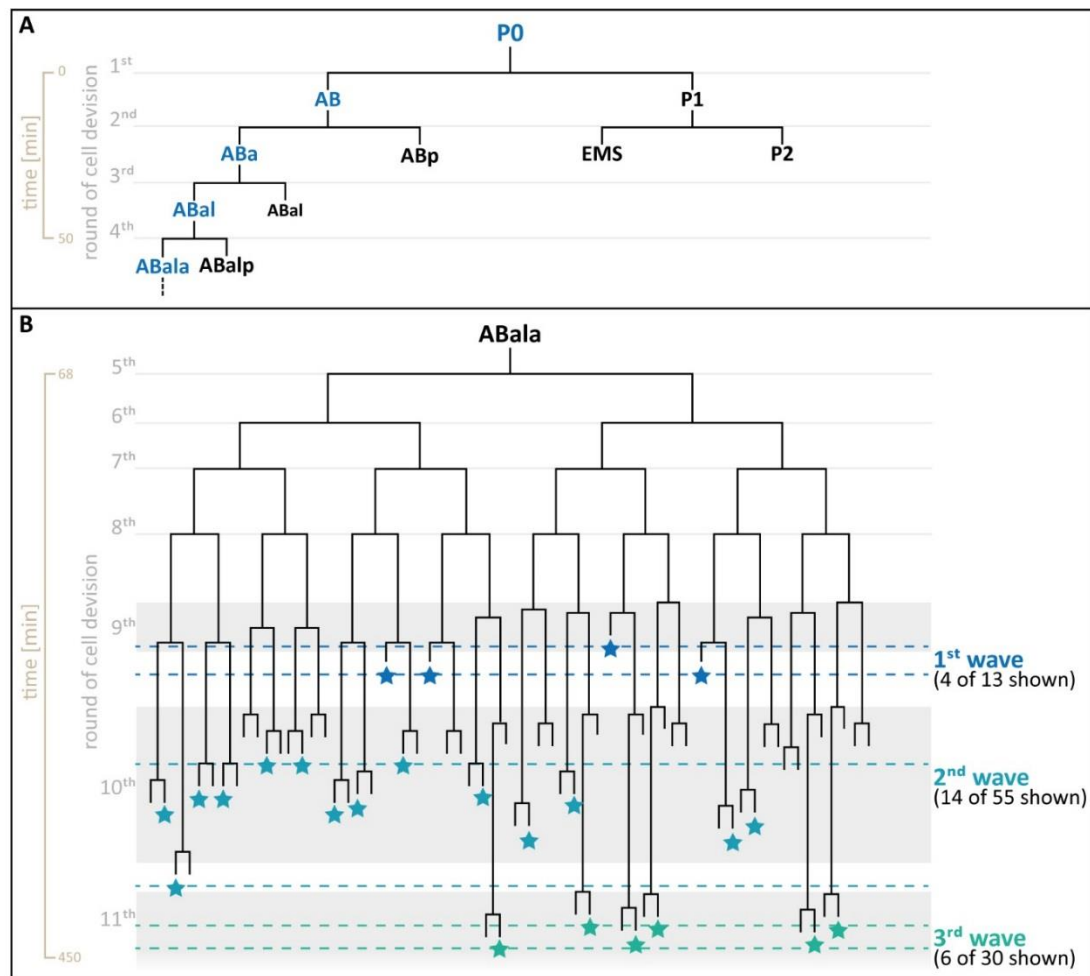


Figure 8. Distinct waves of cell death occur during AB lineage development in *C. elegans*.

A) *C. elegans* lineage tree for ABAla sub-lineage (blue) from P0. **B)** representative sub-lineage tree for ABAla. Brown shows the time from the first cell division, grey the round of cell division. Black represents the lineage tree. Dark blue 1st wave of cell death, light blue 2nd wave of cell deaths and green 3rd wave of cell deaths (adapted from Sulston et al., 1983).

In the examined wildtype *C. elegans* embryos (+/+) cells undergoing cell death round up and form a refractile corpse of ~2.5 μm in diameter which is engulfed by the surrounding cells. These refractile corpses can be identified as “button shaped” in structure with differential-interference-contrast microscopy (Nomaski microscopy) (DIC) (Figure 9 A). For recording of the embryos we have utilized four-dimensional (4D) microscopy (Schnabel *et al.*, 1997, 2006), which adds recording over time (time series) as fourth component in addition to the three axis (X,Y,Z). In wildtype embryos 13 cell deaths were identified as first wave cell deaths within the AB lineage, which are following the *C. elegans* lineage (Sulston *et al.*, 1983). However, in *mir-35* family mutants [*mir-35-41(nDf50) mir-42(nDf49)*], abnormally large cell corpses were identified in addition to the 13 cell deaths, which are also present in wildtype. These large cell corpses have a diameter of about 3.8 μm (1.5-fold), while the prior identified 13 cell deaths do not differ in diameter or engulfment behavior. Further, the abnormally large cell corpses do not show any change in optically visible morphology. The large corpses were never detected before ~180 min within embryonic development, corresponding to the first wave of cell deaths. Upon progression of the embryonic development, these large cell corpses arise more frequent, peaking at ~225 min and again at ~300 min. Large cell corpses are not always engulfed by their surrounding cells. They either persist or were extruded from the embryo into the eggshell. Quantification of these large cell corpses, until ventral enclosure of epidermal cells as determined endpoint at ~330 min (25 °C), results in 4.6 larger corpses per embryo (n=10) (Table 1). To further confirm that the large corpses are dependent on the loss of the *mir-35* family, a rescue experiment was conducted by Ryan Sherrard, utilizing the transgene *nEx1187* that carries *mir-35* alone; the transgenic embryos developed only normal sized cell corpses. It can be concluded that the abnormally large cell-death phenotype is specifically linked to the loss of the *mir-35* family miRNAs (Figure 9 C).

Genotype	# of normal 1 st wave AB cell deaths (n)	# of abnormally large cell corpses (n)
+/+	13 (5)	0 (5)
<i>mir-35-41(nDf50) mir-42 (nDF49)</i>	13 (10)	4.6 +- 0.7 (10)
<i>mir-35-41(nDf50) mir-42 (nDF49); nEx1187[mir-35]</i>	13 (5)	0 (5)
<i>mir-80(nDf53); mir-58.1(n4640); mir-81-82(nDf54)</i>	13 (5)	0 (5)
<i>mir-35-41(nDf50) mir-42 (nDF49); mir-80(nDf53); mir-58.1(n4640); mir-81-82(nDf54)</i>	13 (10)	10.7 +- 0.8 (10)

Table 1: Inappropriate cell death in *mir-35* family mutants is enhanced upon loss of the *mir-58* family

The number of first-wave AB cell deaths was scored per embryo of each genotype. n = 5 or 10, as indicated. The number of abnormally large corpses per embryo was also scored until ventral enclosure. Adapted from Sherrard *et al.*, 2017 available under a Creative Commons License (Attribution 4.0 International).

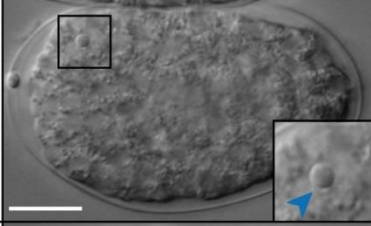
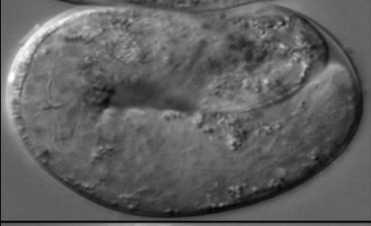
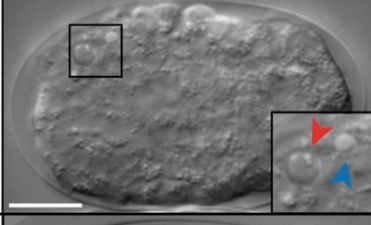
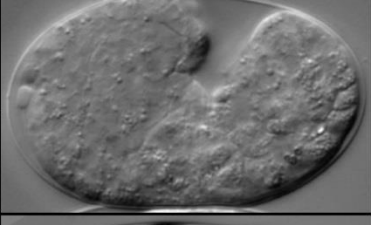
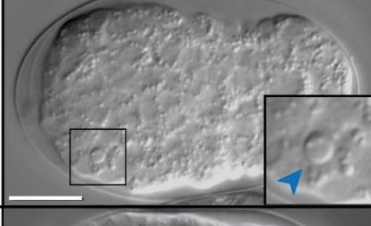
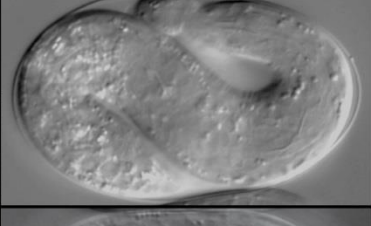
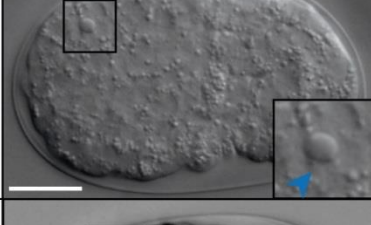

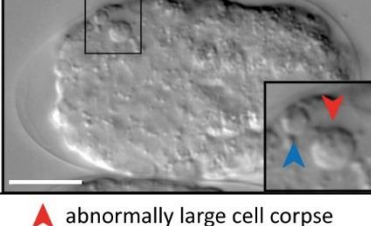
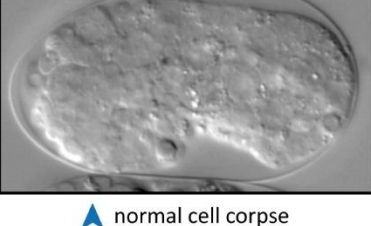


Genotype	Bean stage embryo	Terminal embryonic phenotype	Reference
A +/+			This study
B <i>mir-35-42</i>			This study
C <i>mir-35-42; nEc1187[mir-35]</i>			Sherrard et. al 2007
D <i>mir-80; mir-58.1; mir-81-82</i>			This study
E <i>mir-35-42; mir-80; mir58.1; mir-81-82</i>			Sherrard et. al 2007
 abnormally large cell corpse  normal cell corpse			

Figure 9. Embryos lacking miR-35 family miRNAs exhibit a large cell corpse phenotype.

A–E) Differential interface contrast (DIC) images of embryos of the genotypes **(A)** wildtype, **(B)** *mir-35-41 (nDf50) mir-42(nDf49)*, **(C)** *mir-35-41(nDf50) mir-42(nDf49); nEx1187 [mir-35 (genomic)+sur-5::gfp]*, **(D)** *mir-80(nDf53); mir-58.1(n4640); mir-81-82(nDf54)*, **(E)** and *mir-35-41(nDf50) mir-42(nDf49); mir-80(nDf53); mir-58.1(n4640); mir-81-82(nDf54)* For each panel, a developing embryo (~320-cell stage) is depicted at the left, with insets showing representative cell corpses. Normal cell corpses (blue arrowheads) are present in all genetic backgrounds, but abnormally large cell corpses (orange arrowheads) are also present (see B,E).

The terminal phenotype of each embryo is shown at the right; those in A, C, and D survived to hatching, whereas those in B and E arrested. Scale bars are 10 μ m; references as indicated.

2.2.3 Abnormally large-cell corpse phenotype in miR-35 family mutants is enhanced by additional loss of miR-58 family miRNAs

The *egl-1* 3'UTR also harbors a binding site for the miR-58 family. To investigate the impact on cell death we examined miR-58 family mutants [*mir-80(nDf53)*; *mir-58.1(n4640)*; *mir-81-82(nDf54)*]. As in wild-type embryos normal cell corpses of 2.5 μ m in diameter were detected but no large cell corpses were observed. Further, no embryonic arrest was detected (Figure 9 D), showing that miR-58 family alone is not sufficient to block apoptosis during embryonic development.

A mutant, generated by Nadin Memar, which is lacking the miR-35 and miR-58 family of microRNAs [*mir-35-41(nDf50)* *mir-42(nDf49)*; *mir-80(nDf53)*; *mir-58.1(n4640)*; *mir-81-82(nDf54)*] was examined. As in miR-35 family single mutant, embryos were arresting at around ~500 min after first cell division. The miR-35 miR-58 double-family mutants also display abnormally large cell corpses and normal cell corpses, with a diameter of 3.5 μ m (abnormally large cell corpse) and 2.5 μ m (normal cell corpse) (Figure 9 E). Quantification of abnormally large cell corpses until ventral enclosure (~390 min at 25 °C) shows an average of 10.7 (\pm 0.8) large corpses per embryo which is an 2.3-fold increase compared to miR-35 family single-mutant (4.6 \pm 0.7) and statistically significant (Student's t-test with $P \leq 0.0001$ and $n = 10$) (Sherrard *et al.*, 2017).

2.2.4 The abnormally large-cell death phenotype is dependent on the central apoptotic machinery

The abnormally large cell-death phenotype in the miR-35 family and miR-35 miR-58 family double mutants and the predicted binding sites within the *egl-1* 3'UTR suggest that these additional cell deaths are linked to the central apoptotic pathway. Therefore, I crossed the *mir-35* family mutant into the *egl-1* (*n3330*) deficient mutant as well as into the *ced-3* (*n717*) deficient mutant. Loss of either *egl-1* or *ced-3* blocks nearly all cell deaths in the development of *C. elegans* including the 13 cell deaths present in the first wave of AB lineage cell deaths (Ellis and Horvitz, 1986; Conradt and Horvitz, 1998). For neither of the double mutants, *mir-35* family and *egl-1*, nor the *mir-35* family and *ced-3*, were normal cell deaths or abnormally large cell deaths detected (Table 2). This leads to the conclusion that the miR-35 and miR-58 family act upstream of CED-3 and upstream or in parallel to EGL-1. Interestingly, embryonic lethality was still present and therefore the abnormally large cell deaths are not the cause of embryonic arrest.

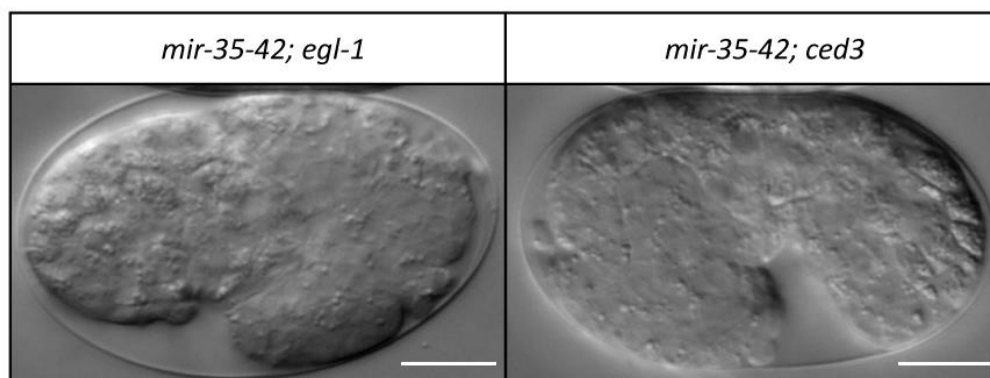


Figure 10. Embryonic lethality is not recovered by blocking apoptosis in miR-35 family mutants.

In both panels DIC images of the terminal phenotype of *mir-35* family mutants also homozygous for *egl-1*(*n3330*) (left) or *ced-3*(*n717*) (right) are shown. Scale bar 10 μm.

Genotype	# of normal 1 st wave AB cell deaths (n)	# of abnormally large cell corpses (n)
+/+	13 (5)	0 (5)
<i>mir-35-41(nDf50) mir-42 (nDF49)</i>	13 (10)	4.6 +- 0.7 (10)
<i>ced-3(n717)</i>	0 (5)	0 (5)
<i>mir-35-41(nDf50) mir-42 (nDF49); ced-3(n717)</i>	0 (5)	0 (5)
<i>egl-1(n3330)</i>	0 (5)	0 (5)
<i>mir-35-41(nDf50) mir-42 (nDF49); egl-1(n3330)</i>	0 (5)	0 (5)

Table 2. miR-35 family miRNAs act upstream or in parallel to *egl-1*.

The number of first-wave AB cell deaths was scored per embryo of each genotype (n = 5). The number of abnormally large corpses per embryo was also scored until ventral enclosure.

Adapted from Sherrard *et al.*, 2017 available under a Creative Commons License (Attribution 4.0 International).

2.2.5 Downregulation of *egl-1* mRNA expression is mediated by the 3'UTR

As shown in 2.2.1 the *egl-1* 3'UTR harbors binding sites for miR-35 and miR-58 family. By fusing the *egl-1* 3'UTR to a reporter construct, the particular role of *egl-1* 3'UTR and its miR binding site can be further elucidated. All generated constructs were integrated via MOS1 transposase-mediated single-copy integration (MosSCI) (Frøkjær-Jensen *et al.*, 2014) and analyzed at four-cell, ~320-cell and ~500-cell stage. Here a *gfp::histone h2b* fusion gene (*gfp::h2b*) under control of the *mai-2* promoter ($P_{mai-2}::gfp::h2b$) was fused to five different combinations of 3'UTRs. The *mai-2* promoter used is ubiquitously expressed in all cells throughout the entire embryonic development and already visible in 2 cell stage.

The first fusion construct contains a *mai-2* 3'UTR, which lacks binding sites for miR-35 and miR-58 and serves as negative control. The generated reporter ($P_{mai-2}::gfp::h2b::mai-2$ 3'UTR) was ubiquitously expressed in all cells throughout the entire embryonic development. In contrast, the *egl-1* 3'UTR fusion construct ($P_{mai-2}::gfp::h2b::egl-1$ 3'UTR) suppresses the expression in four-cell stage as well in ~320-cell stage but shows low level expression at ~500-cell stage. It has been shown that *mir-35-42* expression is lowered after gastrulation and is likely the cause of expression in late stages (Stoeckius *et al.*, 2009; Alvarez-Saavedra and Horvitz, 2010; Isik, Korswagen and Berezikov, 2010; Wu *et al.*, 2010). The reporter was then crossed in the *mir-35* family mutant and has shown the same ubiquitous expression pattern as with the *mai-2* 3'UTR negative control. Further, the *egl-1* 3'UTR reporter construct was mutated at the miR-35 binding site ($P_{mai-2}::gfp::h2b::egl-1^{mut\ mir-35}$ 3'UTR), which shows in wildtype background the same ubiquitous expression pattern as the reporter with *mai-2* 3'UTR in wildtype background and *egl-1* 3'UTR in the *mir-35* family mutant background. Mutating the miR-58 binding site in the *egl-1* 3'UTR within the reporter construct in the wildtype background shows no expression in four-cell and ~320-cell

stage but in the later stage ~500-cell embryo as in the *egl-1* 3'UTR reporter. Finally, the miR-35 and miR-58 binding sites within the *egl-1* 3'UTR were mutated; expression is visible from the four-cell stage throughout the development and seems slightly increased compared with the wildtype *egl-1* 3'UTR in *mir-35* family mutant background as well as in the miR-35 binding site mutated *egl-1* 3'UTR reporter (Figure 11 B).

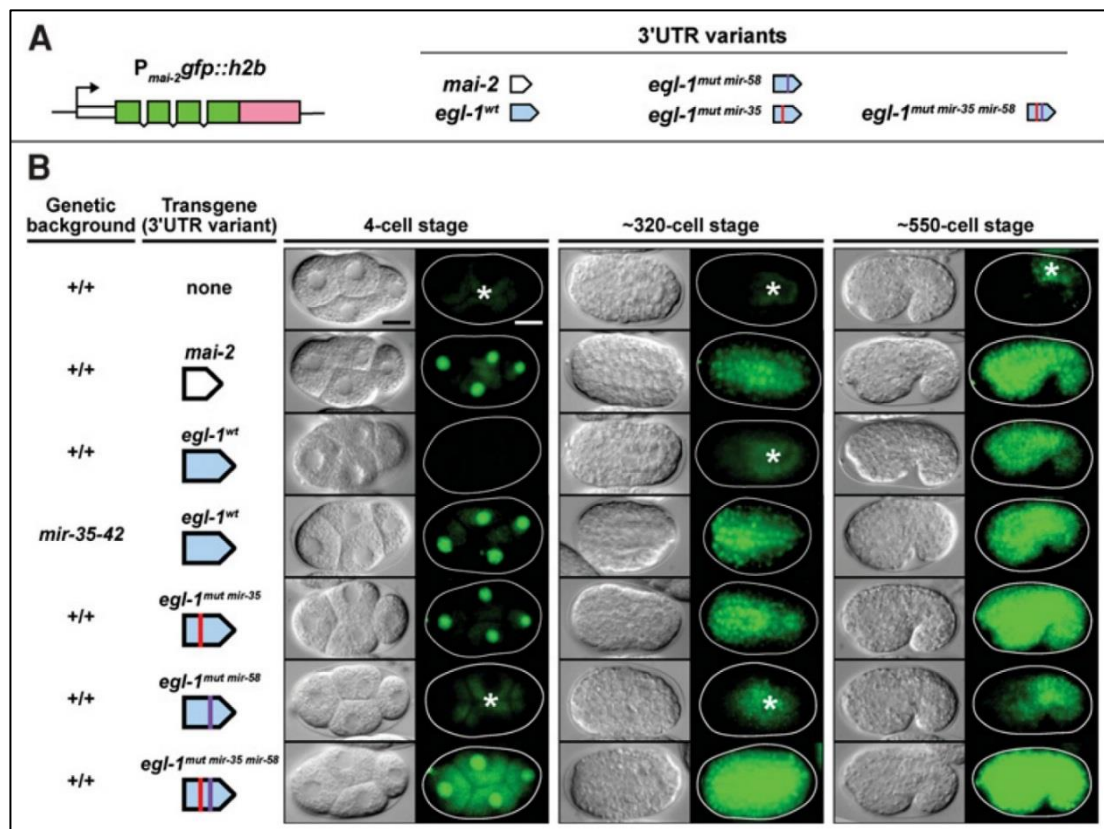


Figure 11. The 3'UTR of *egl-1* is a target of miR-35 and miR-58 family miRNAs *in vivo*

A) A schematic representation of the 3' UTR reporters constructed for this study. **B)** Analysis of single-copy 3' UTR reporter expression during embryogenesis. The genetic background and transgene under investigation are indicated at the left of each image sequence, which shows representative embryos from three developmental stages. For each embryo, a DIC image (left) and GFP image (right) are shown. The following alleles were used: bcSi25 [*P_{mai-2} gfp::h2b::mai-2* 3' UTR], bcSi26 [*P_{mai-2} gfp::h2b::egl-1 wt* 3' UTR], bcSi27 [*P_{mai-2} gfp::h2b::egl-1 mut mir-35* 3' UTR], bcSi46 [*P_{mai-2} gfp::h2b::egl-1 mut mir-35* 3' UTR], and bcSi47 [*P_{mai-2} gfp::h2b::egl-1 mut mir-35 mir-58* 3' UTR]. Asterisks indicate autofluorescence. Transgenic strains were homozygous

for *unc-119(ed3)* and the *cb-unc-119(+)* selection marker. Bars, 10 μ m. Published in and adapted from Sherrard *et al.*, 2017 available under a Creative Commons License (Attribution 4.0 International).

Since miRNAs repress their target mRNA by inhibition of translation or decay of mRNA, determination of which process is responsible for the repression of the *gfp::h2b* reporter is crucial. Here, Ryan Sherrard used single-molecule RNA fluorescent in situ hybridization to directly probe *gfp::h2b* reporter mRNA in transgenic embryos used before. He found that embryos carrying the *mai-2* 3'UTR transgene had a significantly higher copy number of *gfp::h2b* mRNA compared to *egl-1 wt* 3'UTR transgene. Therefore, it is proposed that the miR-35 and miR-58 family miRNAs repress *egl-1* expression by translational inhibition and decay of *egl-1* mRNA.

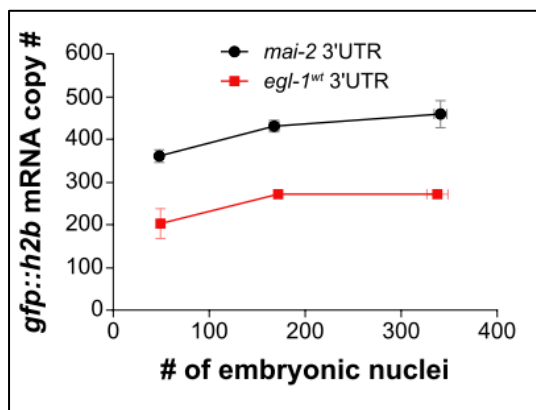


Figure 12. The *egl-1* 3'UTR causes a reduction in mRNA copy number of the P_{mai-2} *gfp::h2b* reporter.

Quantification of *gfp::h2b* mRNA in whole embryos at each stage of interest. n = 3–6 for each data point. Averages are plotted \pm SEM. Adapted from Sherrard *et al.*, 2017 available under a Creative Commons License (Attribution 4.0 International).

2.2.6 Mothers and sisters of normal apoptotic cells die inappropriately in *mir-35* family mutants

Utilizing 4D microscopy I examined five miR-35 family mutants by backwards lineaging to identify all detected abnormally large cell corpses. In some *C. elegans* mutants large cell corpses have been observed before and these deaths have been linked to physically larger cells which undergo programmed cell death. For example as result of early cell death induction or with regards to the asymmetric cell division as a result of a reversed polarity in the dividing mother cell (Hengartner *et al.*, 1992; Sugimoto *et al.*, 2001; Frank *et al.*, 2005). Within these five embryos of the miR-35 family mutant 25 abnormally large cell corpses have been detected; each of these corpses was formed by a precursor cell which does not normally undergo cell death. In detail, all examined large cell deaths could be linked to be either a mother of a cell death (19 out of 25; ~ 76 %) or a sister of a cell death (9 out of 25; ~ 36 %). A special case is the RID progenitor cell (RIDnb – Abalappaap) which is both, a mother and a sister of a cell death and dies inappropriately in 3 out of 5 (60 %) miR-35 family deficient embryos. Mother cells which are dying inappropriately are considered therefore “precocious” cell deaths and sister cells “collateral” cell deaths. As described above, RIDnb a mother and a sister of a cell death dies in 3 out of 5 examined miR-35 family mutants. In miR-35 miR-58 family double knockout mutants RIDnb dies with a penetrance of 80 % compared to miR-35 single mutant (Sherrard *et al.*, 2017). Therefore, miR-58 family knockout enhances miR-35 family dependent formation of abnormally large cell corpses also on single cell level. Interestingly, within the AB lineage I was never able to detect inappropriate cell death in first wave mother cells. Further bilaterally symmetric cell lineages show a similar penetrance upon miR-35 family loss; this was further corroborated by the lineaging experiments done by Ryan Sherrard (Sherrard *et al.*, 2017).

Cell	inappropriate death in mir-35 family mutants			Relation to cell death	Mother / sister of n th wave AB cell death
	This study (n=5)	Sherrard et al., 2017 (n=10)	Sum		
ABplaapapp ¹	4 (80 %)	7 (70 %)	11 (73 %)	mother	3 rd
ABpraapapp ¹	3 (60 %)	6 (60 %)	9 (60 %)	mother	3 rd
ABpraaaapaad ²	3 (60 %)	5 (50 %)	8 (53 %)	sister	2 nd
ABalppppap	3 (60 %)	4 (40 %)	7 (47 %)	mother	2 nd
ABalappaap (RIDnb)	3 (60 %)	4 (40 %)	7 (47 %)	mother / sister	2 nd / 1 st
ABalapapapa ³	2 (40 %)	3 (30 %)	5 (33 %)	mother	3 rd
ABplppaaaap ²	2 (40 %)	3 (30 %)	5 (33 %)	sister	2 nd
ABplppaaaap ⁴	1 (20 %)	3 (30 %)	4 (26 %)	mother	3 rd
ABalapppapa ³	1 (20 %)	2 (20 %)	3 (20 %)	mother	3 rd
ABalppaaaap	1 (20 %)	1 (10 %)	2 (13 %)	sister	1 st
ABarpapaaa	1 (20 %)	1 (10 %)	2 (13 %)	mother	2 nd
MSppaap	1 (20 %)	1 (10 %)	2 (13 %)	mother	-
ABalaappppa ⁵	-	1 (10 %)	1 (7 %)	mother	3 rd
ABalapaaaa	-	1 (10 %)	1 (7 %)	mother	2 nd
ABalapaappa ⁵	-	1 (10 %)	1 (7 %)	mother	3 rd
ABprppaaaap ⁴	-	1 (10 %)	1 (7 %)	mother	3 rd
MSaapaapa	-	1 (10 %)	1 (7 %)	mother	-
Caapa	-	1 (10 %)	1 (7 %)	mother	-

Table 3. Abnormal cell death in *mir-35* family mutants affects mothers and sisters of programmed cell death.

In this study I identified 25 large cell corpses present in five *mir-35-41(nDf50)* *mir-42 (nDf49)* embryos as described in the Materials and Methods. Numerical superscripts 1-5 indicate pairs of bilaterally symmetric cells. For comparison, the large-cell-corpses identified in Sherrard et al. (2017) available under a Creative Commons License (Attribution 4.0 International) were added.

2.3 Analysis of the CRISPR generated *egl-1* 3'UTR variant *bc275*

2.3.1 Generation of *egl-1(bc275)* via CRIPR/Cas9 co-conversion

In 2.2 it was shown that the abnormal large cell deaths in miR-35 family mutants depend on *egl-1*. Mutating the endogenous miR-35 family binding site in the *egl-1* 3'UTR via CRISPR/Cas9 system would exclude side effects arising from the total loss of the miR-35 family during embryonic development, maybe including embryonic lethality. With this motivation the *egl-1* 3'UTR was modified by the CRISPR/Cas9 co-conversion approach (Arribere *et al.*, 2014). The underlying concept of this method is to mutate the gene of interest as well as a selectable marker. This allows for an easy selection of candidate worms directly by their selectable marker, also proving a functional recombination event.

Arribere *et al.* utilized a *dpy-10* loss of function mutation for selection of successfully “converted” worms. Here, within *dpy-10* a C-T substitution is induced, resulting in *dpy-10 (cn64)* variant. Phenotypically this results in dumpy worms, which are shorter and stouter than wildtype animals of the same developmental stage. These worms are therefore easy to spot under a binocular. Selected worms are then backcrossed with wildtype to remove the selectable marker.

Secondly, Ward *et al.*, presented a modified protocol for the co-conversion approach (Ward, 2015). Here temperature sensitive mutant worms which carry the *pha-1* (e2123) mutation are injected and reverted into wildtype via CRISPR/Cas9 induced substitution (Schnabel and Schnabel, 1990; Granato, Schnabel and Schnabel, 1994). Injected worms are grown at 25 °C, at which temperature *pha-1* (e2123) is lethal and only successfully reverted worms are surviving. This approach allows for an easier selection and avoids time consuming phenotypic selection by hand and outcrossing of *dpy-10* (cn64).

The repair template for *egl-1* was designed as single stranded DNA (ssDNA) of 95 bp length complementary to the *egl-1* 3'UTR. The miR-35 family binding site was mutated by switching corresponding base pairs with regards to G/C content of the sequence. Further the repair template does not contain the PAM site of the original sequence to avoid being cut by the Cas9 before or after integration. I have tested both co-CRISPR approaches (see materials and methods).

I injected 234 animals utilizing the *dpy-10* selectable marker, resulting in 57 animals (24 %) showing a dumpy-like phenotype. The co-CRISPR with the *pha-1* based selection results in 4 surviving animals after heat shock out of 200 injected animals (2 %). For both experimental sets it can be concluded the Cas9 was active in cutting the DNA at the selectable marker site. Further, the repair template for the selectable marker was successfully integrated by non-homologous end joining (NHEJ). In direct comparison the *dpy-10* selectable marker has a higher efficiency than *pha-1* selectable marker, according to my findings, 24 % versus 2 % respectively. Within the 57 animals with utilizing the *dpy-10* co-CRISPR approach only two times the *egl-1* 3'UTR was modified. The *pha-1* co-CRISPR did not show any *egl-1* modification (0 out of 4 selectable animals) (Table 4).

Co-CRISPR marker	Injected animals	marker events	<i>egl-1</i> modification
<i>dpy-10</i>	234	57 (24 %)	2 (heterozygous)
<i>pha-1</i>	200	4 (2 %)	0

Table 4: co-CRISPR approach utilizing *dpy-10* and *pha-1* selectable marker for *egl-1* 3'UTR modification.

For each co-CRISPR approach (*dpy-10* or *pha-1* conversion) the injected animals, their respective marker and *egl-1* modification events are listed.

Unfortunately, the donor ssDNA was not incorporated by homologous recombination but instead by non-homologous end joining. The whole donor ssDNA was inserted into *egl-1* 3'UTR next to the induced double strand break the allele was termed *egl-1*(bc275) (Figure 13). Notably, both animals which carry this *egl-1* 3'UTR modification were heterozygous; the repair template was inserted at the exact same position independently. The insertion disrupts the *egl-1* 3'UTR binding site composition by insertion of an additional miR-58 family binding site and leaves the original miR-35-42 family binding site intact. The insertion further disrupts the ALG-1 binding site by one base pair from the beginning and inserts a second mutated ALG-1 binding site, which is mutated within the first 23 base pairs and is missing the last four base pairs.

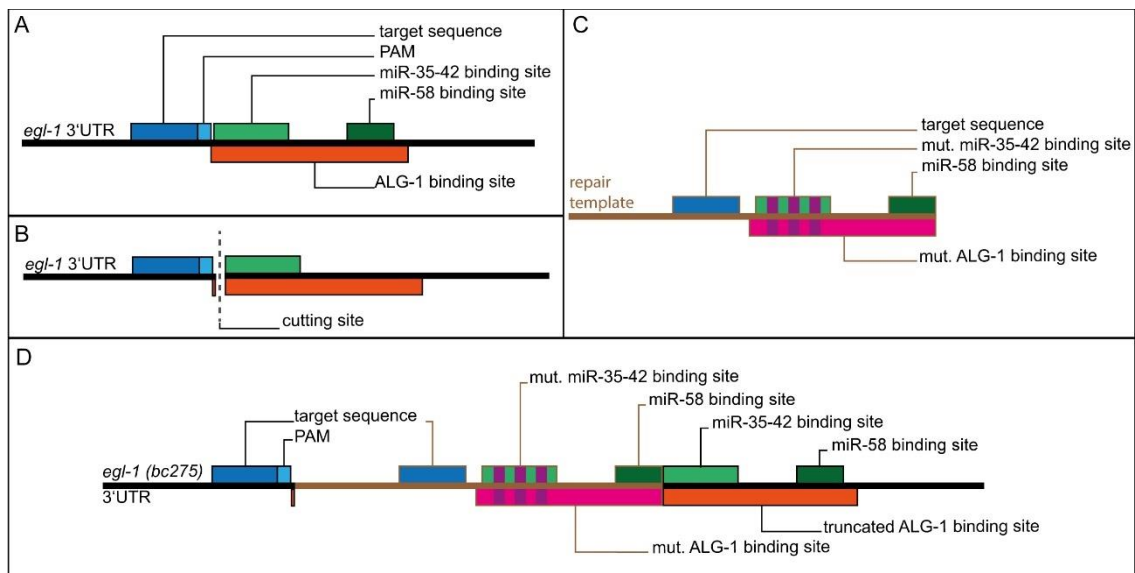


Figure 13: the *egl-1* 3'UTR with inserted repair template after CRISPR/Cas9 induced NHEJ.

A) The *egl-1* 3'UTR (152 bp) with important binding-sites indicated; miR-35-42 binding site (light green), miR-58 binding site (dark green), ALG-1 binding site (orange), and the sgRNA targeting sequence (dark blue) and PAM site (light blue). **B)** The cutting site within the *egl-1* 3'UTR truncates the first four base pairs of the ALG-1 binding site. **C)** The repair template is based on the *egl-1* 3'UTR and spans 59 bp up- and downstream of the defined cutting site leaving the miR-58 binding site intact but truncates the ALG-1 binding site by four base pairs. Further the miR-35-42 family binding site was mutated by base pair flipping, which also mutates the

overlapping ALG-1 binding site by 23 base pairs. The miR-58 family binding site remains unchanged by design. The CRISPR/Cas9 targeting sequence is intact but the PAM site was mutated to prevent cutting of the repair template. **D)** The *egl-1* 3'UTR was cut as shown in (B) and the repair template (C) was integrated via non-homologous end joining, resulting in the *egl-1* (*bc275*) variant. The variant contains two CRISPR/Cas9 targeting sequences for the same sgRNA but only one PAM site. Further two intact miR-58 family binding sites a miR-35-42 family binding site and a mutated miR-35-42 binding site. The ALG-1 (pink) was mutated together with the miR-35-42 family binding site and is further truncated by 23 bp. The second ALG-1 binding site (orange) is truncated by the first four base pairs.

2.3.2 The gene variant *egl-1* (*bc275*) is causing a general block in apoptosis

The insertion mutation *egl-1* (*bc275*) causes a change in the predicted folding of the *egl-1* 3'UTR (Figure 14). While for the wildtype *egl-1* 3'UTR the positional entropy is overall low, reflecting a rather stable 3'UTR folding, the *egl-1* (*bc275*) mutation increases the overall entropy and results in unstable *egl-1* 3'UTR folding with a different structure compared to the *egl-1* wildtype 3'UTR. These changes in stability and folding may impact the binding capabilities of interaction partners within the *egl-1* (*bc275*) variant.

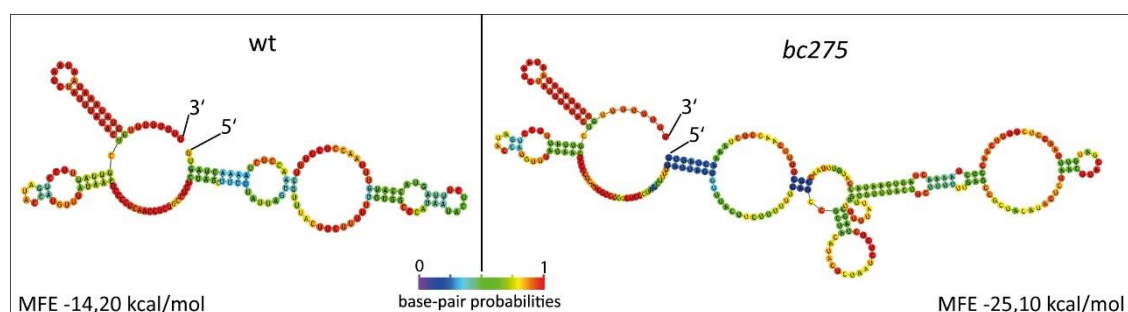


Figure 14: The donor ssDNA integration changes the predicted *egl-1* 3'UTR folding.

A) predicted folding of the wildtype *egl-1* 3'UTR . **B)** Predicted folding of the *egl-1* (*bc275*) 3'UTR.

The colors represent the base-pair probabilities entropy. The minimum free energy (MFE) is

increased in the *egl-1 (bc275)* variant over wildtype (wt), indicating a higher thermodynamic stability. Prediction from RNAfold web server (<http://rna.tbi.univie.ac.at>) (Gruber *et al.*, 2008; Lorenz *et al.*, 2011)

C. elegans carrying the *egl-1(bc275)* allele display a cell-death-defective phenotype, meaning that during embryonic development no cell deaths are observable. Further, developing embryos always have cells which are extruded and persist throughout the development (n=10) (Figure 15).

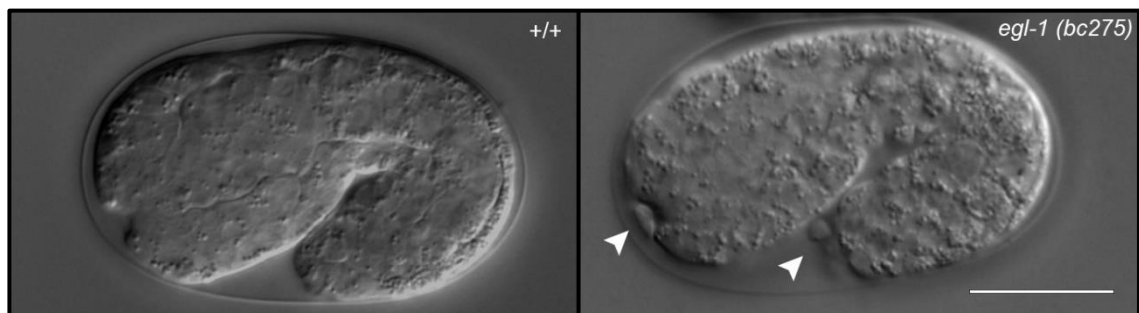


Figure 15: Extruded cells in embryos carrying the *egl-1 (bc275)* variant.

In *egl-1 (bc275)* (right) animals no cell death is visible during embryonic development (n=10); extruded cells (white arrows) are pushed into the eggshell. In wildtype animals (left) cells are not extruded from the embryo (scale bar 10 μ m).

For further examination of *egl-1 (bc275)* on the cell death execution, extra cells in the pharynx were counted together with Barbara Conradt. Within the anterior pharynx 16 cells normally undergo programmed cell death; in the event of defects in the programmed cell death pathway these cells persist as “undead” extra cells. This method provides a very sensitive quantitative assay at single-animal level for detecting defects in cell death execution (Schwartz, 2007). Wildtype animals only show rare cases of extra cells with an average of one extra cell per 10 animals (Conradt and Horvitz, 1998; Schwartz, 2007). Mutations which block the programmed cell death pathway in *C. elegans* (*egl-1 (lf)* / *ced-3 (lf)* / *ced-4 (lf)* / *ced-9 (gf)*) are causing on average 11-12

extra cells within the anterior pharynx (Hengartner *et al.*, 1992). Within the *egl-1 (bc275)* variant we find on average 11.4 (n=17) extra cells in the anterior pharynx. These results indicate that *egl-1 (bc275)* is blocking apoptosis in *C. elegans* development.

Genotype	mean	SEM	range	Reference
+/+	0.1	0.1	0-1	(Schwartz, 2007; Breckenridge, Kang and Xue, 2009)
<i>egl-1 (bc275)</i>	11.4	0.312	9-13	This study

Table 5: Extra cells in the anterior pharynx.

Within wildtype (N2) animals on average 0.1 extra cells are present (see References). *egl-1 (bc275)* animals display 11,4 extra cells on average with a standard error of mean (SEM) of 0.312. On single animal level the extra cells range from 9 to 13.

2.3.3 Translation of *egl-1* mRNA is blocked by *bc275* 3'UTR mutation

I have shown that the *egl-1* (*bc275*) variant causes a general block in apoptosis during *C. elegans* development. Since *egl-1* (*bc275*) is an insertion mutation within the *egl-1* 3'UTR, the assumption that the mRNA is misregulated is an obvious guess. To verify if *egl-1* mRNA is present at all, single molecule fluorescent in situ hybridization (smRNA FISH) was used to determine the mRNA status.

Here I'm utilizing the *unc-3* reporter ($P_{unc-3}unc-3::gfp$) which is expressed in the RID lineage (Wang *et al.*, 2015). The mother cell RIDnb (ABalappaap) and the two daughter cells, the RID neuron (ABalappaapa) and the dying sister (ABalappaapp), can be identified by GFP expression (mother and dying daughter) and their position in the developing embryo.

In wildtype animals *egl-1* mRNA is largely absent within the RIDnb, at the time point close to division. Here, I measured an average of 0.6 *egl-1* mRNA copies (n=5; range 0-1). After division I measured a concentration of 1.4 *egl-1* mRNA copy numbers (n=5; range 1-2) in the RID neuron. In the dying sister cell, I measured 10.4 copies of *egl-1* mRNA (n=5; range 8-12). In total the *egl-1* mRNA copy numbers in wildtype animals within the RID lineage are not significantly different from what has been shown before (Sherrard *et al.*, 2017). Interestingly, when looking at the *egl-1* (*bc275*) variant within the RID lineage, the RIDnb already shows a significantly higher ($P < 0.01$) amount of 2.75 *egl-1* mRNA copy numbers compared to wildtype. First, this shows clearly that *egl-1* mRNA is present within this variant. Secondly, the *egl-1* mRNA copy number should be sufficient to trigger cell death within in this lineage. The threshold of around 2 *egl-1* mRNA copies was shown before (Sherrard *et al.*, 2017). After division of the RIDnb, I measured 1.5 *egl-1* (*bc275*) mRNA copies (n=5; range 1-2) in the RID neuron, which is not significantly different from what I measured within the wildtype (1.4 *egl-1* mRNA copies), indicating that the *egl-1* mRNA turnover in the RID neuron is independent from the *egl-1* 3'UTR. Interestingly, within the dying sister I measured a significantly higher

($P < 0.01$) *egl-1* (*bc275*) mRNA copy number of 15.5 ($n=5$; range 14-17) compared to the 10.4 *egl-1* mRNA copy number in wildtype. Further, in the dying daughter of the RIDnb, Sherrard et al. could not find a significant difference between wildtype and the miR-35 family, miR-58 family and the double knockout, with regards to *egl-1* mRNA copy number. The *egl-1* mRNA copy number present in this variant is significantly higher ($P < 0.01$). This indicates that there is potentially another way of *egl-1* mRNA regulation within the 3'UTR of *egl-1*.

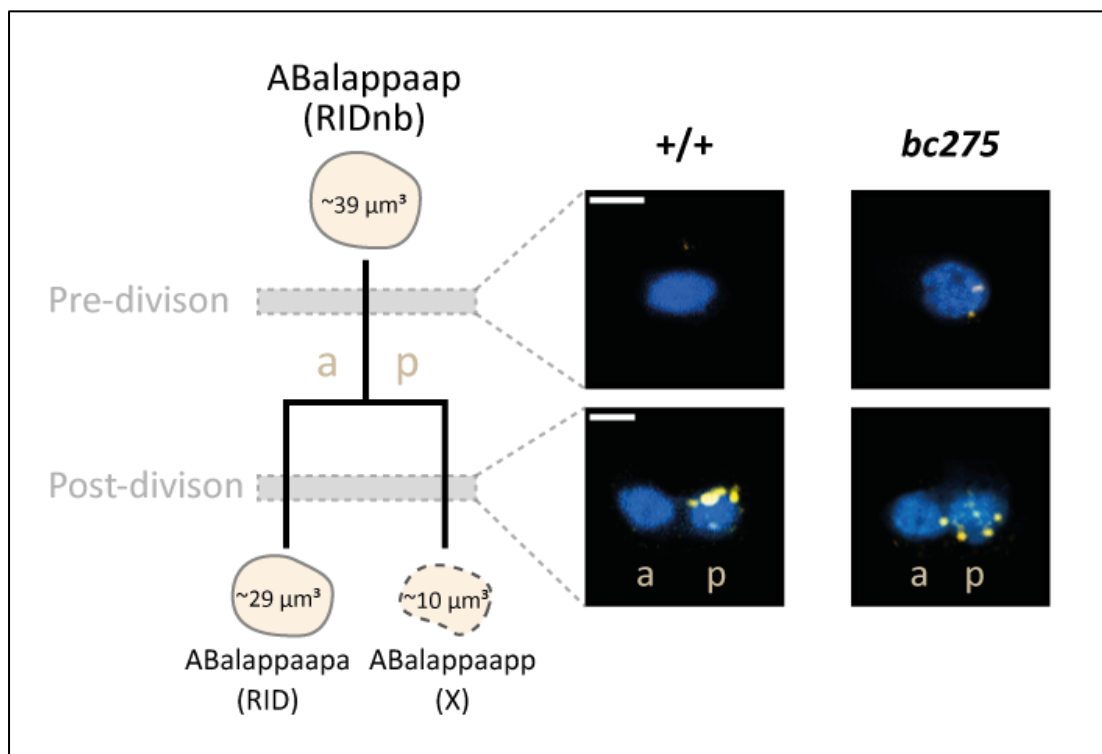


Figure 16: Representative smRNA FISH images of *egl-1* mRNA in the RID lineage

smRNA FISH analysis of the RID lineage in wildtype (+/+) and *egl-1* (*bc275*) variant animals. The RID lineage is illustrated at the left, indicating the pre- and post-division as time point of measurements. As indicated on the right, representative states of the RIDnb and its daughters are shown for each time point wildtype and *egl-1 bc275* variant background. The nuclei of anterior (a) surviving and posterior (p) dying daughter and the RIDnb stained with DAPI, smRNA FISH labeled *egl-1* mRNA is shown in yellow. (Maximum intensity z-projection with Gaussian Blur, radius = 1.5, DAPI signal from neighboring cells was removed). Scale bars, 2 μm .

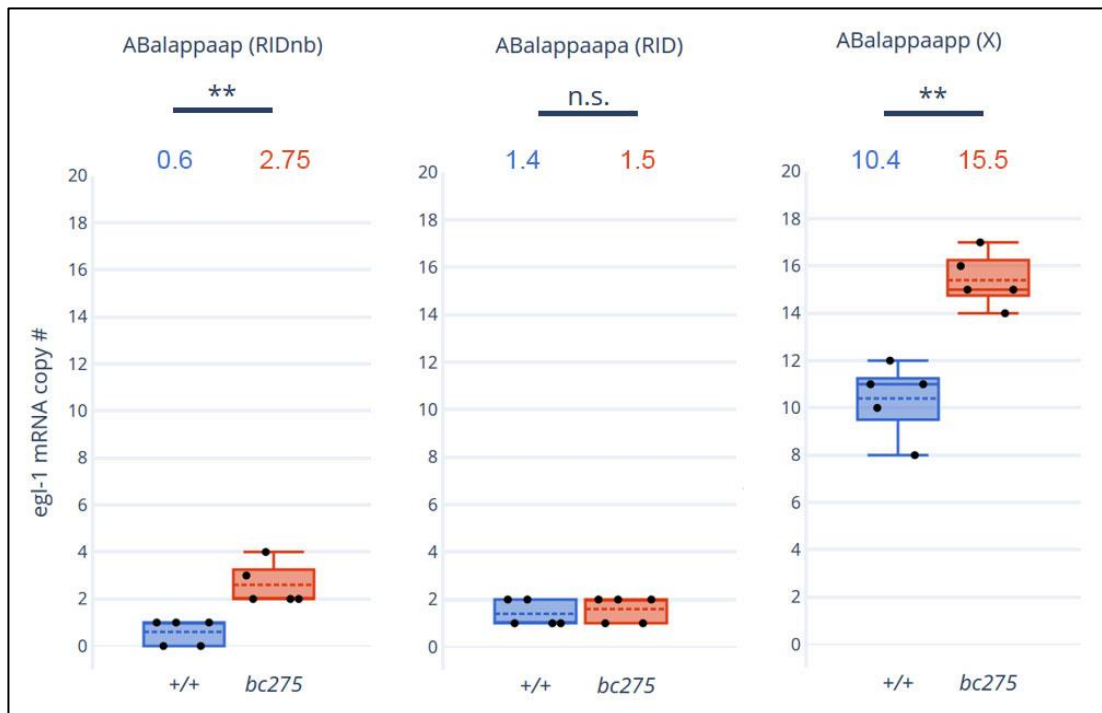


Figure 17: Quantification of *egl-1* mRNA copy number in the RID lineage.

The RIDnb (ABalappaap) and its two daughters ABalappaapa (RID) and ABalappaapp (X) for each of the indicated genotypes. The mean value is given above each dataset. (n.s.) not significant; (**) $P < 0.01$ via Mann-Whitney test. The wildtype (+/+) (blue) genotype is compared to the *bc275* allele (red). All strains analyzed carried the extrachromosomal array *xdEx1091* [*P_{unc-3} unc-3::gfp+P_{sur-5} rfp*]; $n=5$.

2.4 Fluorescent CIRSPR/Cas9 mediated tracing of distinct DNA sequences

2.4.1 Generation of an inactivated dCas9::GFP::GFP fusion construct

One of the main features of the CRISPR/Cas system, used in laboratories, is the targeting of specific custom designed DNA sequences and double strand break induction at a defined position, while being easily programmable and cost effective. For utilizing the CRISPR/Cas system for visualizing distinct sequences it is necessary to inactivate the endonuclease activity by introducing mutations to the Cas9 gene. Therefore I mutated a *C. elegans* codon-optimized version of the Cas9 gene (Ward, 2015) at the amino acid positions 10 and 840, replacing asparagine with alanine and histidine with alanine, respectively. These mutations inactivate the RuvC1 (D10) and HNH (H840) endonuclease sites, leading to the nuclease deficient form termed dCas9. I fused the dCas9 gene to two GFP genes via a glycerin-serine linker and two SV40 nuclear targeting sequences (NLS). Here I used two GFPs to achieve a higher signal intensity. Further, both GFPs are codon-optimized for *C. elegans* but differ in their codon usage to avoid recombination or excision of the gene or parts of it. The *dCas9::gfp::gfp* construct was fused to an *mai-2* promoter under control of the *mai-2* 3'UTR. Between the second GFP gene and the 3'UTR a second SV40 NLS has been additionally integrated. For simplicity I will call the construct *dCas9::gfp::gfp* fusion. I have shown in chapter 2.2.5 that *Pmai-2* promoter and *mai-2* 3'UTR led to expression within all cells beginning already in the two cell stage of embryonic development, ensuring a stable expression throughout the entire lifecycle.

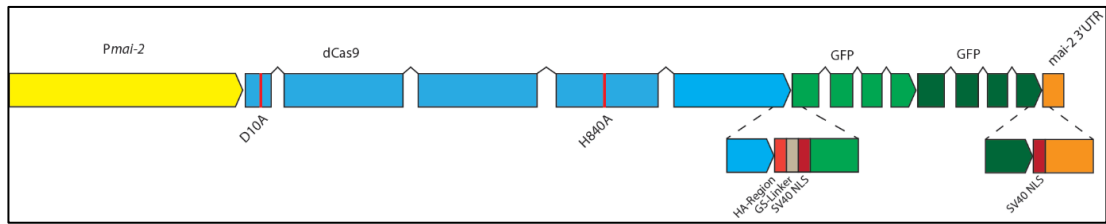


Figure 18. Labeling targeted DNA sequence with a dCas9::GFP::GFP.

The *mai-2* promoter (yellow) and the *mai-2* 3'UTR (orange) drive expression of the dCas9::GFP::GFP fusion. The dCas9 (blue) with inactivated RuvC1 (D10A) and HNH (H840A) (red bars) is fused via a HA-region, glycine-serine linker and an NLS to two GFPs which differ in their codon usage (dark green / light green).

2.4.2 The dCas9::GFP::GFP fusion is embryonic lethal in *C. elegans*.

The *dCas9::gfp::gfp* fusion construct was subsequently cloned into the integration vector used for *mos1* mediated single copy integration (MosSCI) (see materials and methods). In total I injected 1621 worms to achieve integration of the construct on linkage group I or IV; to exclude toxic effects I varied concentrations of the construct from 25 ng/μl to 1ng/μl. I was able to achieve 17 % to 29 % of *unc-119* rescued F1 animals after injection, but these drop dramatically in the F2 generation and in later generations the injected construct was lost (see Table 6). By decreasing the concentration, it was possible to achieve a higher rate of *unc-119* rescued animals. Nevertheless, it was neither possible to get an integration into the linkage-group I or IV nor a stable extra-chromosomal array. This indicates a toxicity of the *dCas9::gfp::gfp* fusion construct. This could be further validated when looking at the *Pmyo-3::mCherry* co-injection marker (5 ng/μl) which was expressed in dead embryos and is used to identify extrachromosomal arrays in injected worms.

Linkage group	Concentration of	Worms injected	<i>unc-119</i> rescued F1 (Array)	<i>unc-119</i> rescued F2 (Array)	<i>unc-119</i> rescued F ^x (Array) silenced or lost
I	25 ng/μl	154	28 (18 %)	0	0
	5 ng/μl	132	31 (23 %)	0	0
	2,5 ng/μl	354	79 (22 %)	0	0
	1 ng/μl	420	121 (28 %)	5	5
IV	25 ng/μl	54	11 (20 %)	0	0
	5 ng/μl	100	18 (17 %)	0	0
	2,5 ng/μl	162	38 (23 %)	1	1
	1 ng/μl	245	72 (29 %)	3	3

Table 6: *C. elegans* injected for *mos1*-mediated single copy integration of the *dCas9::gfp::gfp* fusion construct.

2.4.3 Telomere targeting sgRNAs

Besides the toxicity upon injection of the construct, I could detect GFP expression within all *unc-119* rescued F2 generation animals. Therefore, I wanted to determine if the *dCas9::gfp::gfp* fusion is located within the nucleus and if it can be directed to a target site. First, I wanted to target the repetitive sequence of the telomeres in *C. elegans* (TTAGGC). Since NGG is the defined PAM site of the used dCas9 (see introduction), the target sequence is also defined (see Figure 19). The pattern of telomere locations has already been studied in *C. elegans* using different methodology. For example, it has been shown by DNA FISH experiments that there are on average 18.6 detectable telomeric foci per cell (Ferreira *et al.*, 2013). Telomeres can be therefore used as validation for a functional methodology.

I injected the designed telomere targeting sgRNA (50 ng/μl) along with the *dCas9::gfp::gfp* fusion construct (1 ng/μl) into a strain which has a single copy integration of an EMR-1 tagged with mCherry (bqSi142 [*emr-1p::emr-1::mCherry* + *unc-119(+)*] II). Since EMR-1 is located at the inner membrane of the nucleus it is the ideal marker to show localization of the *dCas9::gfp::gfp* fusion construct.

After injection, the F1 generation of L4 stage animals were screened for expression of GFP. As expected from the toxicity of the *dCas9::gfp::gfp* fusion construct, only a few animals show GFP expression. Further, all examined animals showed a mosaic pattern of expression. The *dCas9::gfp::gfp* fusion was correctly located in the nucleus but it was not possible to detect distinct foci of dCas9::GFP::GFP (see Figure 19). The dCas9::GFP::GFP fusion was not directed to the telomeres, instead an accumulation of the dCas9::GFP::GFP fusion was detectable within the nucleoli of the nucleus. This suggests that either the dCas9 is not fully functional or that the level of expressed sgRNAs was too low. The latter option would be supported by observation that the dCas9 can non-specifically interact with ribosomal RNA or nuclear proteins (Chen *et al.*, 2016).

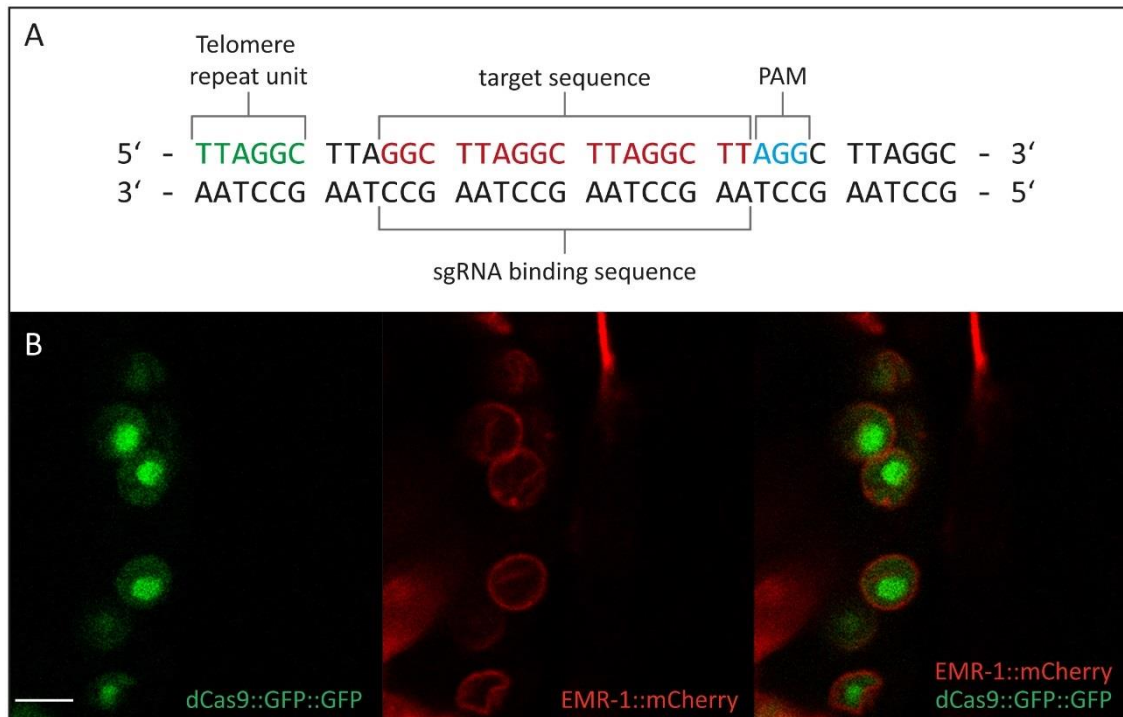


Figure 19. Targeting telomeric repeats with dCas9::GFP::GFP/sgRNA complex.

A) Telomere targeting sgRNA design; in *C. elegans* the telomeres are comprised of the repeating element 5' – TTAGGC – 3' (green); the target sequence (red) which is part of the sgRNA located next to PAM site (blue). **B)** Images taken of the tail cells of a L4 animal. The left panel shows dCas9::GFP::GFP fusion construct; the center panel the EMR-1::mCherry; and the right panel the merge of both. The dCas9::GFP::GFP fusion is located within the nucleus of the cell but accumulates within the nucleoli. No directed localization of the dCas9::GFP::GFP fusion construct to the telomeric repeats is detectable, Scale bars, 2 μ m.

2.5 *In silico* identification of chromatin modifiers with potential impact on *egl-1*

2.5.1 Hypothesis: chromatin modifiers open the *egl-1* locus prior to first wave of cell death

The majority of apoptosis during embryonic development of *C. elegans* occurs in three distinct waves (see Figure 8, chapter 1.2.2 and 2.2.2). The first wave of cell deaths arises at around the 9th round of cell division, here the embryo has around 150 cells. Before this time point cell death is not observed. This is also reflected within the *egl-1* expression pattern (see Figure 20). I have shown before that in the AB lineage mother cells of the first wave of cell deaths are not showing any inappropriate cell death in miR-35 family mutants. One could speculate that this is supported by chromatin modifications within the genome, which open the *egl-1* locus for transcription prior to the first wave of cell death. The logical chain can therefore be described as follows: A chromatin modifying gene needs to be expressed (or expression is suppressed), resulting in a change of the chromatin composition at *egl-1* locus supporting its expression. With regards to timing this would lead to the 8th round of cell division, since EGL-1 is already present in the mother cell of first wave cell deaths. The chromatin modifying gene, therefore, needs to change its expression pattern at or most probably before the 7th round of cell division. This regulation could be reflected within the expression level of the chromatin modifying gene.

2.5.2 Identification of chromatin modifiers by the *C. elegans* gene ontology

For identification of all chromatin modifying genes I used a candidate list which is based on *C. elegans* gene ontology (GO) terms and their association to genes. Since the gene ontology does not harbor enough hierarchic and synonym information to directly select all chromatin modifying genes, a keyword-based full-text search was conducted over all gene ontology terms and their descriptive texts. All GO Term IDs which are associated with one keyword of a defined list were selected (see Table 7), resulting in a list of 912 associated GO terms. Interestingly, some GO terms of these do not contain an associated gene, indicating that curation of the gene ontology is still work in progress (and will probably always be). From this set of GO terms, all associated WormBase IDs (gene ID) were selected from the database, leading to a pool of 625 unique genes. In a second step all genes containing the keywords in their description were selected, regardless of their gene ontology group, resulting in 196 associated unique genes. Both sets of genes were merged into pool of 693 candidate genes. Ideally, the keyword associated genes would already be part of the GO term generated list, indicating again that data curation is an ongoing process. The candidate list was further reduced from 693 to 585 candidate genes since these candidate genes do not have an expression TPM (transcripts per million) above zero.

keyword	keyword associated GO terms	GO terms associated genes	keyword associated genes
nucleosome	47	105	27
nucleosome modifi	0	0	0
histone	430	420	141
histone modifi	127	210	3
chromatin	260	407	85
chromatin factor	9	99	12
chromatin modification	9	95	0
chromatin modifier	0	0	0
chromatin modifi	12	95	1
chromatin remod	18	38	11
candidate genes (unique)		1469 (625)	280 (196)

unique candidates with TPM > 0	585
------------------------------------------	------------

Table 7. Collection of candidate genes via Gene Ontology by keyword

For each keyword (1st column) the associated GO terms were selected (2nd column), which either have the keyword in their name or description; of these GO terms all connected genes were selected, merged and duplicates were removed (3rd column). Second for each keyword all genes with associated genes were selected (4th column), which either have the keyword in their name or description. Third, the total candidate genes of GO term pool and directly associated genes are merged and duplicates are removed. Finally, candidate genes which have a TPM (transcripts per million) above zero have been selected for further analysis.

2.5.3 Quality control of candidate genes

For quality control I've generated a list of 72 known *C. elegans* chromatin modifiers based on wormbook and a review (Cui and Han, 2007; Ahringer and Gasser, 2018). This list also contains homologs from yeast, *Drosophila melanogaster*, and/or mammals. Interestingly, through mapping of homologs from other organisms to *C. elegans*, the wormbook list from 2007 contains genes which have just recently been shown to also play a role in *C. elegans*. The keyword-based approach I presented before, was able to identify 67 of 72 chromatin factors. The remaining five genes could not be identified for three reasons. First, their description is not containing any of the keywords and second they are not associated with an appropriate gene ontology group. This has often one of the following reasons, first the gene is not actively studied within the *C. elegans* community, so metadata enrichment is poor for these genes. Second, the gene is studied, but results published have not been incorporated into the *C. elegans* gene ontology. Third, the gene has a known homolog in another organism which is known for chromatin interactions, but the biological function has not been shown in *C. elegans* and is therefore not represented in the *C. elegans* gene ontology (Table 8).

identifier	gene name	human ortholog	last publication	notes
WBGene00009924	<i>cfp-1</i>	CXXC1	2019	gene ontology is missing data from publications
WBGene00003034	<i>lin-49</i>	BRD1	2010	gene ontology is missing data from publications
WBGene00010677	<i>gtbp-1</i>	G3BP1	2018	only one publication but not related to chromatin; only listed as human ortholog
WBGene00011887	<i>set-17</i>	PRDM7	2018	only one chromatin related publication; results not represented in gene ontology
WBGene00020314	T07E3.3	GSTK1	2006	<i>C. elegans</i> : protein with unknown functions; Human: chromatin remodeling study in SynMuv pathway; listed as human ortholog

Table 8: Genes not identified based on keyword-based search within the *C. elegans* gene ontology.

2.5.4 Expression data analysis of chromatin associated candidate genes

When looking at the time course of the *egl-1* transcription (Figure 20) (Hashimshony *et al.*, 2015), it becomes apparent that *egl-1* peaks between the 7th and 8th data point of the expression dataset. This peak coincides with the first wave of cell death (Figure 20 A). Following the hypothesis presented above, transcription changes of the chromatin modifier should take place at the 6th data point (referred to as point of interest; POI), when the embryo has around 110 cells. I have defined the idealized expression pattern of a candidate gene which is upregulated as follows (Figure 20 B):

- 1st, the TPM at the POI is the highest TPM in the time course
- 2nd, the TPM at the POI is above average
- 3rd, the TPM left (POI - 1) and right (POI + 1) from the POI is at least 5 % (10 % or 20 %) below the POI value
- 4th, the TPM two steps left (POI - 2) and right (POI + 2) from the POI is at least 5 % (10 % or 20 %) below the POI value
- 5th, the transcription TPM on the left (POI - 1) is below and on the right (POI + 1) above the POI
- 6th, the TPM on two steps left (POI - 2) is below and two steps right (POI + 2) is above the POI

For candidate genes which are getting downregulated the data is calculated conversely (Figure 20 C); calculations are described in detail in chapter 4.7. For each condition which is true, the candidate gene is ranked higher. The ranking within the 3rd and 4th condition is also influenced by the percentage; here higher percentages are ranked higher. One has to keep in mind that the 3rd or the 4th condition is not valid if the 5th or the 6th condition is true and vice versa. But rising (or declining) TPMs are preferred over steady TPMs.

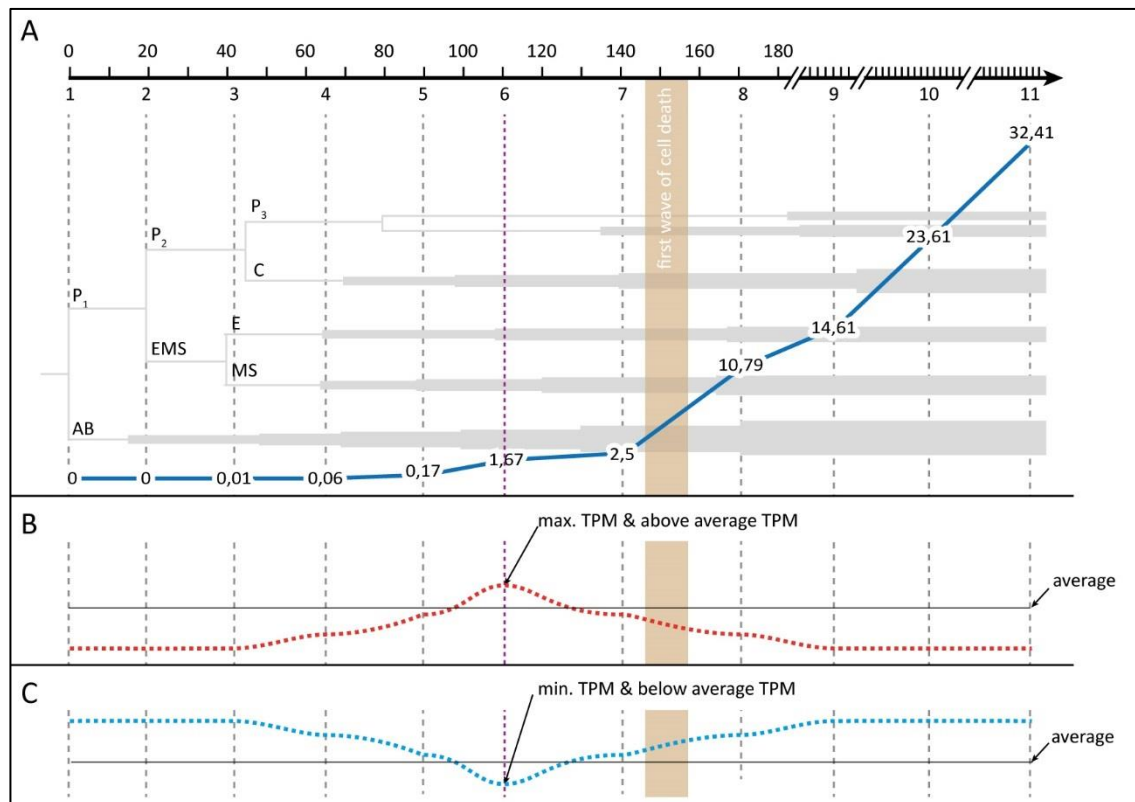


Figure 20: *egl-1* expression level (TPM) during embryonic development

A) The *egl-1* expression (blue line) peaks for the first time at around the first wave of cell death, where the embryo has round 150 cells. In the expression dataset used here this corresponds to the 8th data point. Here I am focusing on the 6th data point as the mostly likely time point were changes in the chromatin modifying enzymes could be detected. B) The red dotted line shows an idealized expression pattern of a generic chromatin modifier which is upregulated shortly before the first wave of cell death. It peaks at the 6th data point and the transcripts per million (TPM) are above average and additionally represent the maximum of TPM measured (global maximum). Further, TPM levels, left and right of the 6th data point, are decreasing. (C) The idealized downregulated chromatin modifier would look like the horizontal mirrored upregulated chromatin modifier (blue dotted line). The x-axis represents time in minutes after the first AB division with expression data time points shown below. Expression data for *egl-1* from (Hashimshony *et al.*, 2015)

For the 585 candidate genes it is possible to reach a maximum score of eight points with different combinations. In total 55 candidates reached this maximum. 34 of these are upregulated and 20 are downregulated. Of the candidates which reach seven points 26 are upregulated and 24 are downregulated. For 29 of these candidates, the global maximum was below 5 % so they are not considered as up- or downregulated here. Of the 36 candidates which reach six points 10 are upregulated and 7 are downregulated, while 19 are neither upregulated nor downregulated. As for candidates with five points 10 were upregulated, 18 were downregulated and 10 are below the 5 % border. This is the first time more candidates are downregulated than upregulated. This is also the case for four, three and two points candidates. Of the four-points candidates, 27 are upregulated and 100 are downregulated, while 90 are below the 5 % border. Within the candidate group which reaches three points only two are upregulated, while 13 are downregulated and three are showing a peak which is below 5 %. Finally, within the 2 points group 10 are upregulated, 45 are down regulated and 88 are below the 5 % border (Table 9).

Points	Upregulated	Downregulated	Global maximum < 5 %	Total
8	34	20	0	54
7	26	24	29	79
6	10	7	19	36
5	10	18	10	38
4	27	100	90	217
3	2	13	3	18
2	10	45	88	143
Total	119	227	239	585

Table 9: Candidate genes are mapped against expression pattern and sorted in categories.

Candidate genes were assigned points bases and their expression pattern as described in 0. While in total 119 of the 585 candidates are upregulated, 227 genes are downregulated, and 239 genes have a global maximum which is below 5 %.

For candidates reaching the maximum score I wanted to determine how good these fit to the idealized expression pattern as defined in Figure 20 B and C. Therefore, I plotted ten representatives of up- and down-regulated genes and their expression data during embryonic development (see Figure 21). As the graphical representation shows the self-developed Python application produces the expected outcome.

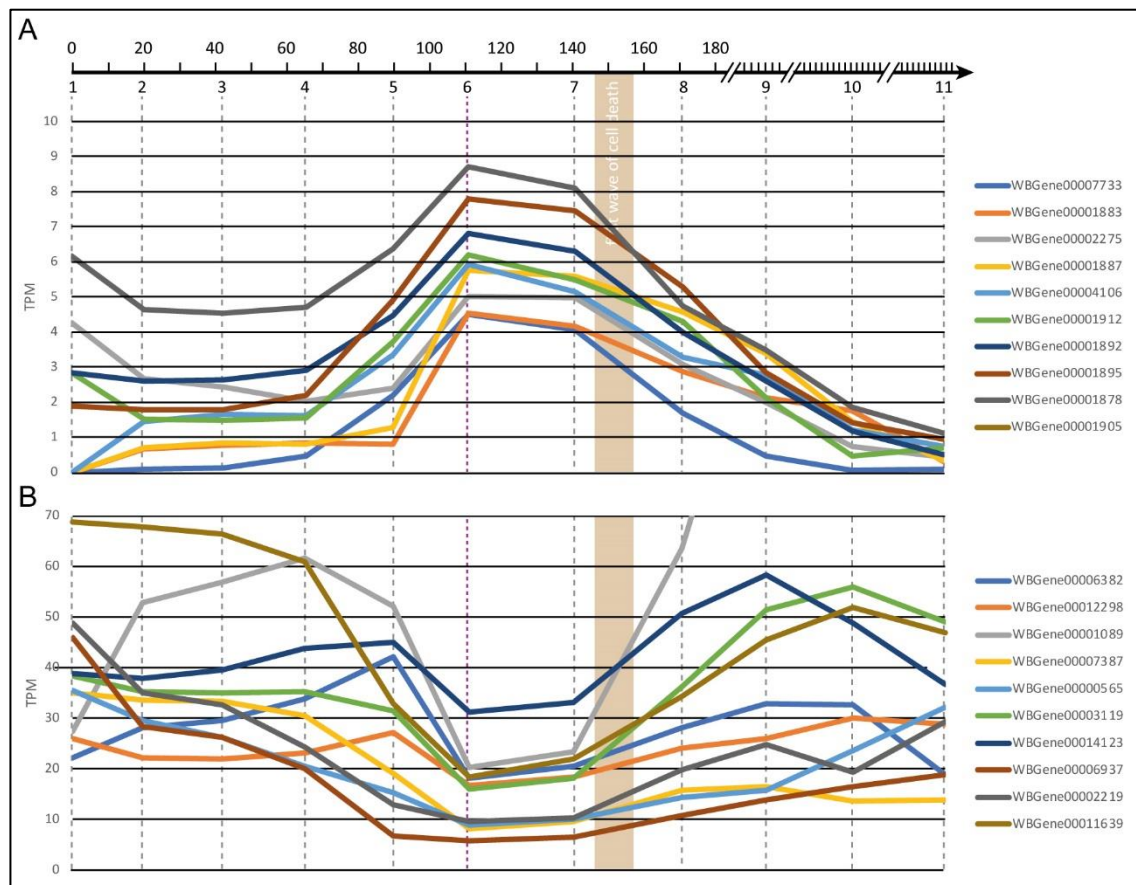


Figure 21: Ten representative candidate genes with best fit (eight points) to the defined expression pattern.

For both diagrams the y-axis represents the transcripts per million (TPM) and the x axis represents time in minutes after the first AB division with expression data time points shown below. Upregulated (A) and downregulated (B) potential chromatin modifiers which fit the defined idealized expression pattern (compare Figure 20). Expression data from (Hashimshony *et al.*, 2015)

2.5.5 Overrepresentation analysis of candidate groups associated with GO term “programmed cell death”

As described before I have generated the candidate gene dataset by their association with the defined keywords and their expression pattern (see 2.5.2). But I was also curious if GO terms are enriched within this dataset. Therefore, a GO overrepresentation test over the biological processes powered by the PANTHER classification system was done (see Materials and Methods) (Mi et al., 2019) (<http://pantherdb.org>). One has to keep in mind, that this analysis can only show already known and annotated biological processes.

In a first analysis I specifically looked at the overrepresentation of the GO term “programmed cell death” (GO:0012501) overall ranked candidate groups (Table 10). For the highest ranked (eight points) candidate group with 54 candidates, seven were associated with the targeted GO term. With expected 0.23 genes associated with this GO term this results in a 30.24-fold enrichment. For the seven-points target group which comprises of 79 candidates, seven genes were associated the GO term “programmed cell death” (expected 0.37) resulting in a 19.01-fold enrichment. The six points group hold 36 candidate genes of which three were associated with “programmed cell death” (expected 0.17) which results in a 17.82-fold enrichment. In the five-points target group with 38 candidates the enrichment increases to 20.1-fold with four genes associated with “programmed cell death” while 0.20 were expected. The group of 217 genes rated with four points, is the biggest candidate group and holds 29 genes associated with the GO term “programmed cell death”, while 1.13 were expected, resulting in a 25.64-fold enrichment. Within the three points group two of the 18 candidates were associated with “programmed cell death”, 0.11 were expected. Interestingly, for this group the fold enrichment is not significant since the group of 18 candidates is too small leading to a P value smaller than 0.05 when Bonferroni-corrected. For the target group holding two

points in the ranking 12 candidates of 143 were associated with “programmed cell death”, while 0.74 were expected, leading to 16.18-fold enrichment.

Interestingly, the GO term programmed cell death is in general enriched overall candidate groups, indicating a strong connection of programmed cell death with chromatin modifications within the *C. elegans* gene ontology. Although the highest ranked group also has the highest association with the GO term, there is no declining trend visible overall groups.

		<i>C. elegans</i> (REF)	dataset			
points	# of candidates	#	#	expected	fold enrichment	raw P value
8	54	106	7	0.23	30.24	4.61E-09
7	79	106	7	0.37	19.01	1.27E-07
6	36	106	3	0.17	17.82	6.78E-04
5	38	106	4	0.20	20.1	5.34E-05
4	217	106	29	1.13	25.64	7.24E-30
3	18	106	2	0.11	(18.33) ¹	(1.00) ¹
2	143	106	12	0.74	16.18	3.61E-11

Table 10: Overrepresentation analysis for GO term programmed cell death (GO:0012501) overall candidate groups.

Shown are the points for each candidate group (1st column) and the number of candidates within this group (2nd column). Within the *C. elegans* reference 105 genes were associated with the GO term programmed cell death (GO:0012501) (3rd column) and were analyzed against genes with the same association in the candidate groups (4th column). The expected value shows the number of expected genes in this category based on the reference list (5th column). The fold enrichment represents the genes observed in the dataset list over the expected (6th column). The raw P value shows the probability that the number of genes observed in this category occurred by chance (randomly) (7th column). ¹ is not significant, Bonferroni-corrected for $P < 0.05$. (Data calculated via PANTHER classification system (Mi *et al.*, 2019); <http://pantherdb.org>; PANTHER14.1; released 20200407; settings in Materials and Methods.)

I further submitted the candidate list to FunSet for enrichment visualization (Hale, Thapa and Ghera, 2019) (Figure 22). Here I selected all genes with five or more points (206 genes in total), since I consider these as best fitting to my hypothesis based on their expression patterns. These genes were clustered into 30 individual groups by FunSet. These clusters represent significantly enriched terms associated in semantically similar groups. These clusters are then ordered on a 2D plot by their respective aggregate information content capturing the semantic similarity between terms. This makes it possible to uncover biological process features which are shared within the candidate list. In this clustered representation, chromatin and especially methylation related GO terms are enriched, which is expected since the analysis is based on the chromatin associated gene list. Further 4 of 13 enriched GO term groups are associated with apoptosis (3) or with developmental processes (1) (Table 11). Further pointing in the direction that genes which are known chromatin factors are also playing a role in apoptotic and developmental processes or even have a dual role.

GO ID	GO term name
GO: 0006915	apoptotic process
GO: 0050793	regulation of developmental process
GO: 0010942	positive regulation of cell death
GO: 0097193	intrinsic apoptotic signaling pathway

Table 11: Apoptotic and developmental process related enriched GO terms within the candidate list

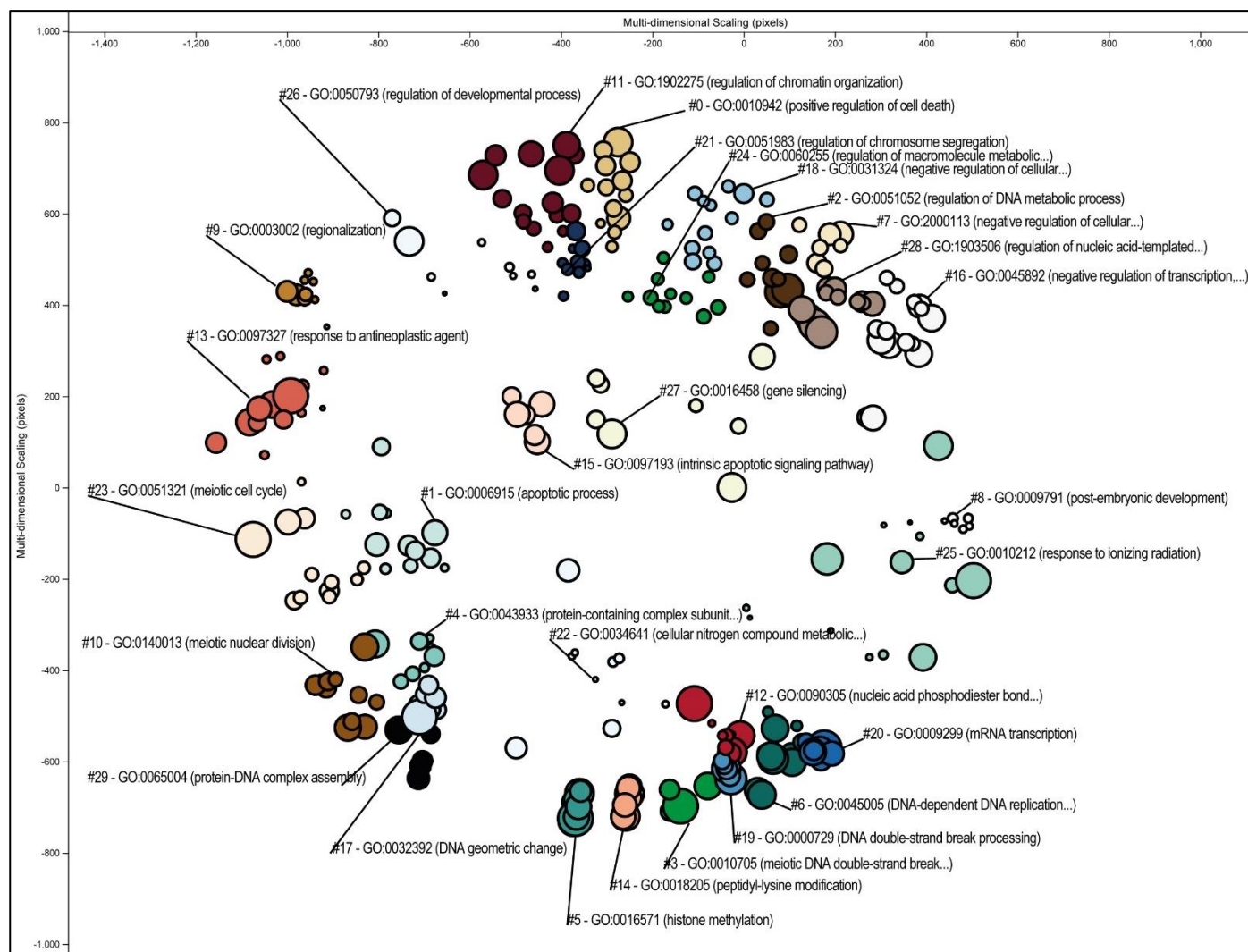


Figure 22: GO enrichment analysis visualization

The data set is comprised of 206 candidate genes with more than five points in the expression data scoring. Each colored bubble represents an enriched gene ontology term. Bubbles of the same color represent clusters of associated gene ontology terms. Picture adapted from FunSet (Hale, Thapa and Gheri, 2019). FunSet settings: January 2019 GO Ontology; Organism: *C. elegans*; Namespace: Biological Process; FDR threshold 0.05; Clusters 30.

2.5.6 Overrepresentation analysis of the highest ranked candidate group

Looking closer at the highest ranked 54 candidates (eight points), four of the top five enriched GO term groups can be directly associated with chromatin modifications (Table 12). Which is expected since the candidate genes were selected as chromatin modifiers. Interestingly, when looking at the candidate genes associated with these GO term groups none of them is currently also associated with programmed cell death.

	<i>C. elegans</i> (REF)	dataset			
GO biological process complete	#	#	expected	fold enrichment	raw P value
protein ADP-ribosylation (GO:0006471)	6	2	.01	> 100	1.31E-04
histone H3-K36 methylation (GO:0010452)	6	2	.01	> 100	1.31E-04
response to X-ray (GO:0010165)	11	2	.02	82.48	3.63E-04
protein deacetylation (GO:0006476)	17	3	.04	80.06	1.10E-05
histone lysine demethylation (GO:0070076)	14	2	.03	64.81	5.56E-04

Table 12: Top five enriched GO terms of candidate genes with ideal expression pattern.

Shown are the highest enriched GO terms (and their ID) based on the candidates having the ideal defined expression pattern (1st column). Here the mapped genes within the *C. elegans* reference (2nd column) are compared with the mapped genes in the dataset (3rd column). The expected value shows the number of expected genes in this category based on the reference list (4th column). The fold enrichment represents the genes observed in the dataset list over the expected (5th column). The raw P value shows the probability that the number of genes observed in this category occurred by chance (randomly) (6th column). (Data calculated via PANTHER classification system (Mi *et al.*, 2019); <http://pantherdb.org>; PANTHER14.1; released 20190711; settings in Materials and Methods.)

Further within these candidate list, eight GO term groups which are associated with cell death or apoptosis are enriched (Table 13). While genes within GO term group cell death are 29.92-fold enriched it increases within the groups programmed cell death and positive regulation of cell death to 30.77-fold and 31.22-fold, respectively. Within the GO term group programmed cell death, apoptotic process is 35.61-fold enriched, regulation of programmed cell death 30.87-fold and positive regulation of programmed cell death 43.12-fold, which is the highest enrichment within this dataset. Below the apoptotic process GO term group the two subgroups positive regulation of apoptotic process and regulation of apoptotic process are enriched 39.95-fold and 29.79-fold, respectively. Interestingly the two highest enrichments can be observed within GO term groups associated with a positive regulation of either apoptotic process (39.95-fold) and programmed cell death (43.12-fold).

	<i>C. elegans</i> (REF)	dataset			
GO biological process complete	#	#	expected	fold enrichment	raw P value
cell death (GO:0008219)	120	7	0.24	29.92	5.29E-09
└programmed cell death (GO:0012501)	103	7	0.23	30.77	4.12E-09
└└apoptotic process (GO:0006915)	89	7	0.2	35.61	1.58E-09
└└└positive regulation of apoptotic process (GO:0043065)	34	3	0.08	39.95	7.36E-05
└└└regulation of apoptotic process (GO:0042981)	76	5	0.17	29.79	9.3E-07
└└└└positive regulation of programmed cell death (GO:0043068)	42	4	0.09	43.12	3.13E-06
└└└└regulation of programmed cell death (GO:0043067)	88	6	0.19	30.87	5.66E-08
└└└└└positive regulation of cell death (GO:0010942)	58	4	0.13	31.22	1.04E-05

Table 13: Section of the GO term enrichment analysis of candidate genes with ideal expression pattern.

Shown are enriched GO terms (and their ID) with their relation to the GO term “cell death”, based on the candidates having the ideal defined expression pattern (1st column). The tree of the *C. elegans* gene ontology reference is indicated in orange. The mapped genes within the

C. elegans reference gene ontology (2nd column) are compared with the mapped genes in the dataset (3rd column). The expected value shows the number of expected genes in this category based on the reference list (4th column). The Fold Enrichment represents the genes observed in the dataset list over the expected (5th column). The raw P value shows the probability that the number of genes observed in this category occurred by chance (randomly) (6th column). (Data calculated via PANTHER classification system (Mi *et al.*, 2019); <http://pantherdb.org>; PANTHER14.1; released 20190711; settings in Materials and Methods.)

Looking closer at the GO term “apoptotic process” seven individual genes of the 43 candidates are associated. These genes are known to play a role in cell death and its regulation (Table 14). Notably, these seven genes are not directly related to chromatin modifications but were identified by being associated with one or more associated keywords. This is related to the keyword-based analysis which can contain false positive genes. Interestingly, HLH-3 as part of the HLH-2/HLH-3 heterodimer has been shown to be a direct, cell-type specific activator of *egl-1* transcription (Thellmann, Hatzold and Conradt, 2003).

WBGeneID	gene name	interaction with <i>egl-1</i>	reference
WBGene00006937	<i>wah-1</i>	Released by <i>egl-1</i>	(Wang <i>et al.</i> , 2002)
WBGene00003119	<i>mac-1</i>	None; binds to <i>ced-4</i> and prevents cell death	(Wu <i>et al.</i> , 1999)
WBGene00001950	<i>hlh-3</i>	Activates <i>egl-1</i> transcription	(Thellmann, Hatzold and Conradt, 2003)
WBGene00001089	<i>dre-1</i>	None; Parallel to EGL-1; inactivates CED-9	(Chiorazzi <i>et al.</i> , 2013)
WBGene00000565	<i>cnt-1</i>	None; cleaved by CED-3; inactivates the AKT cell-survival pathway.	(Nakagawa, Sullivan and Xue, 2014)
WBGene00000420	<i>ced-6</i>	None; involved in engulfment of apoptotic cells	(Liu and Hengartner, 1998)
WBGene00000146	<i>ape-1</i>	Not directly; APE-1 inhibits CEP-1, CEP-1 activates EGL-1	(Lettre <i>et al.</i> , 2004; O’Neil and Rose, 2006)

Table 14: Cell death associated candidate genes identified via GO enrichment analysis

Table shows genes which are associated with the enriched GO terms shown in Table 13. These genes were identified via chromatin associated keyword search in the *C. elegans* gene ontology.

Their expression pattern matches the hypothetic ideal pattern for *egl-1* activation (Figure 20 B).

For HLH-3 an direct activation of *egl-1* has been shown (Theilmann, Hatzold and Conradt, 2003).

3 Discussion

3.1 Post-transcriptional regulation of *egl-1* by microRNAs

3.1.1 Repression of *egl-1* by miR-35 and miR-58 family microRNAs is enhanced by cooperative effect

The data collected by R. Sherrard, N. Memar and me, provides evidence that the predicted binding sites for the miR-35 family and miR-58 family of microRNAs within the *egl-1* 3'UTR are indeed targeted *in vivo*. Confirming the proposal from Wu et al., based on *in vitro* studies, miR-35 and miR-58 family miRNAs can repress *egl-1* mRNA (Wu et al., 2010). Binding of both miR families to their target site leads to suppression of *egl-1* expression. The absence of the miR-35 family leads to precocious and collateral death in cell-death lineages. These deaths are phenotypically visible as large-cell-corpse phenotype in examined cell lineages of miR-35 family mutants. The lack of suppression is further confirmed by derepression of the *P_{mai-2}gfp::h2b::egl-1* 3'UTR reporter I generated. Although, the large-cell-corpse phenotype is not present in the absence of the miR-58 family, I could show that the *P_{mai-2}gfp::h2b::egl-1* 3'UTR reporter is also derepressed when miR-58 family is depleted. The miR-35 family and miR-58 family binding sites within the *egl-1* 3'UTR are 33 bp apart from each other, which is within the distance of 13-35 bp where microRNA cooperativity was shown (Grimson et al., 2007; Saetrom et al., 2007). MicroRNA cooperativity is defined as the synergistic effect of two or more microRNAs on their target mRNA, in this case an enhanced suppression of expression. The results show that indeed a cooperative effect enhances the observed large-cell-corpse phenotype upon loss of both miR families. Further, the derepression of the *P_{mai-2}gfp::h2b::egl-1* 3'UTR reporter was enhanced compared to the single knockout of either the miR-35 family or mir-58 family microRNAs. R. Sherrard additionally reports significantly higher *egl-1* mRNA levels in RIDnb in the absence of both mRNA families (Sherrard et al., 2017). This finding is confirmed by studies in the mammalian system, which support microRNA cooperativity resulting in an enhanced silencing capability (Broderick et al., 2011). The RIDnb also undergoes precocious cell

death in 47 % of the examined embryos upon loss of miR-35 family microRNAs. This post-transcriptional control is critical to avoid precocious death of the RIDnb, by keeping the *egl-1* mRNA level below a threshold sufficient to trigger death. A genetic model for this cell lineage is proposed in Figure 23.

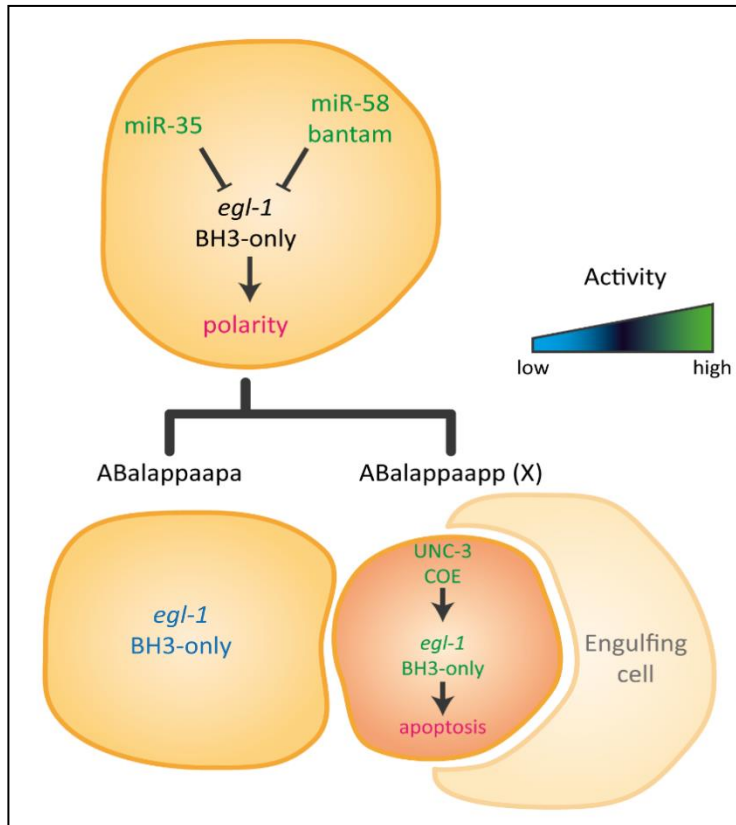


Figure 23. Genetic model of miR-35 and miR-58 family dependent control on *egl-1* BH3-only in the RID lineage.

Within the mother cell of the RID neuron (RIDnb/ABalappaap), the miR-35 family and miR-58 family of microRNAs dampen *egl-1* expression post-transcriptionally, keeping the *egl-1* mRNA level above the threshold necessary for cellular polarization and below the threshold to trigger *egl-1* mediated cell death. After division *egl-1* activity is low in the surviving daughter (RID / ABalappaapa). By contrast in the dying sister cell (ABalappaapp) *egl-1* expression is boosted by the transcription factor UNC-3/COE above the threshold needed to trigger apoptosis. Figure adapted from Sherrard *et al.*, 2017 available under a Creative Commons License (Attribution 4.0 International).

3.1.2 miR-35 family and miR-58 family microRNA activity equilibrates lineage-specific differences in *egl-1* transcription

During *C. elegans* development exactly 131 cells which are programmed to die have a high level of *egl-1* activity. Within these lineages, leading to cell death, *egl-1* activity is tightly controlled by transacting-factors (reviewed in Conradt et al. (2016)). These factors have a varying, yet distinct impact on *egl-1* transcription depending on the cell lineage. These lineage specific differences in *egl-1* transcription, not only change the factor which activates *egl-1* transcription, but also the timepoint and the degree of expression initiation. This may explain the varying penetrance among the cell death lineages upon depletion of *mir-35* family and *mir-35 mir-58* double-family. This is further supported by a comparable penetrance observed in bilaterally symmetric lineages, in which *egl-1* expression is controlled by similar transacting-factors. In the RID lineage, the transcription factor UNC-3 is required for terminal differentiation of the RID neuron. By contrast in the dying sister cell, UNC-3 induces cell death by direct interaction with the *egl-1* locus (Wang *et al.*, 2015). In the RID neuron EGL-1 is therefore needed in a concentration allowing differentiation but not crossing a threshold to induce inappropriate cell death. This duality of EGL-1 is also present within the NSM lineage. Here EGL-1 is already active in the NSM neuroblast and required for cellular polarization and segregation of apoptotic potential into the dying daughter cell prior to cell division (Chakraborty *et al.*, 2015; Lambie and Conradt, 2016). From this point, the miR-35 family as well as the miR-58 family adjust lineage-specific differences in *egl-1* expression. This ensures in the end a level of EGL-1 sufficient to fulfill its non-apoptotic function but preventing inappropriate cell death.

3.1.3 The miR-35 and miR-58 families of microRNAs buffer *egl-1* expression during embryonic development

The *mir-35* and the *mir-58* family of microRNAs are expressed with different abundance during embryonic development. The *mir-35* family is highly expressed in oocytes and during early and mid-stage embryos, while the *mir-58* family is expressed with increasing levels from mid-stage-embryogenesis throughout the whole development. While these expression patterns partially overlap, the impact they have on *egl-1* expression differs. Based on the *egl-1* 3'UTR reporter assays I presented, the repression by the miR-35 family is stronger compared to the miR-58 family. This has been shown before by *egl-1* 3'UTR luciferase-based *in vitro* assays and has been later further confirmed by smRNA-FISH experiments (Wu *et al.*, 2010; Sherrard *et al.*, 2017). Further, I did not observe any inappropriate cell death within *mir-58* family mutants. While the miR-58 family represses *egl-1* expression, it does not take over the role of the miR-35 family in later embryonic stages with regards to inappropriate cell death prevention. With regards to the weaker miR-58 mediated repression the binding site position within the *egl-1* 3'UTR may also play a crucial role. It has been shown in the mammalian system that microRNA binding-sites which are close to the beginning or the end of the 3'UTR tend to be more active compared to binding-sites located centrally (Grimson *et al.*, 2007). Interestingly, 3'UTR shortening can lead to an enhanced repression of centrally located microRNA binding-sites and has been shown to promote tumor growth *in vivo* by triggering proto-oncogene expression in *cis* (Masamha *et al.*, 2014; Xia *et al.*, 2014). The waves of cell death observed in *C. elegans* embryonic development (see Figure 8) are affected differently by loss of the examined miR families. Lineaging results I presented earlier show that mothers of the AB cell lineage are affected to 0 % in first wave in *mir-35* family mutants. In later stages, penetrance is increasing to 1.8 % and 8.0 % in the second and third wave of cell deaths, respectively (Sherrard *et al.*, 2017). While there is a clear effect of the miR-35 family loss in the

developing embryo, it has a low penetrance. This suggests that other *egl-1* regulatory mechanisms play a more essential role in cell death initiation and that these are overall sufficient in the context of normal embryonic development. This would indicate, in this context, that the role of both examined miR families have is the creation of an equilibrium between the non-apoptotic function of *egl-1* and cell death initiation in certain cell lineages. Interestingly, it has been shown recently that the miR-35 family is also buffering MAPK activity (MPK-1 regulates *egl-1* expression in response to DNA damage) by repressing *ndk-1* in germ cells and may represent a failsafe mechanism to genotoxic stress (Tran *et al.*, 2019). The resistance of first wave mother cells to miR-35 family loss, that I presented earlier, is striking. Here, one could speculate that the cytoplasmatic volume may play a role. Cells within the first wave of cell death have a larger volume than cells in later stages. With shrinking volume, the potential of interaction increases and may lead to a lower resistance with regards to varying *egl-1* expression levels. While there is an additional miR-230 binding site in the *egl-1* 3'UTR it is unlikely to have an impact on embryonic development. First, it is predicted to bind with a lower probability compared to the miR-35 and miR-58 family. Secondly, miR-230 expression has only been shown from L1 to the L3-larval stage (Martinez *et al.*, 2008). Nevertheless, it seems worth mentioning that *mir-230* has been shown to be downregulated together with *mir-35*, *mir-37*, *mir-38* and *mir-40* in fine particulate matter toxicity tests in *C. elegans* (Wu *et al.*, 2017). Another point could be downstream factors of EGL-1 (CED-4 and CED-3 but not CED-9) maybe limiting cell death induction leading to an ostensible resistance. Also, EGL-1 may be regulated on a protein level or the *egl-1* locus may not be accessible to its *trans*-acting factors at this timepoint. The underlying hypothesis is that at least before the 9th round of cell division (beginning of the first wave of cell deaths) the *egl-1* locus is in a closed heterochromatic state and only becomes accessible in mothers of cells programmed to die mediated by chromatin modifiers. This hypothesis is the foundation for the *in silico* analysis I performed in chapter 2.5 and is discussed in chapter 3.4.

3.2 Analysis of the CRISPR/Cas9 generated *egl-1* 3'UTR variant *bc275*

3.2.1 Modification of the *egl-1* locus by CRISPR/Cas9

I have shown that the *egl-1* locus can be targeted by CRISPR/Cas9 and have compared two different selection marker strategies. While selection via the *pha-1* re-conversion is the more elegant approach, it was also, for me, the least efficient. I was able to achieve a conversion rate 2 %, compared to that of 24 % using the *dpy-10* marker. While there are approaches available which foster a more efficient selection, introduction of additional elements (e.g., fluorescent marker or drug resistance) or leaving a scar (e.g. Lox-P site) in the target gene is required. The different efficiencies between the two markers may be caused by the chromatin structure at the marker site and recent publications show that the Cas9 binding to DNA is influenced for example by the nucleosome density in close proximity to the target site (Hinz, Laughery and Wyrick, 2015; Horlbeck *et al.*, 2016). The genomic composition could also be the reason why the *egl-1* conversion happens with such a low frequency. Using the co-conversion approach, one has to keep in mind that *egl-1* might be converted more frequently but not together with the marker and is therefore missed. In an ideal setup, the selection marker and the targeted gene should have the same conversion rate. Since this cannot be determined in advance, the current approach is to use a low efficient marker to reduce the selection effort needed (Dickinson and Goldstein, 2016). Secondly, the donor ssDNA was not incorporated by homologous recombination but by non-homologous end-joining. This has been shown previously for dsDNA repair templates and for plasmid repair templates. For the latter a frequency of 5-10 % has been reported (Dickinson and Goldstein, 2016). Therefore, in future experiments it is recommended to inactivate or downregulate nonhomologous end joining, for example by RNAi, which has been reported to increase homologous recombination efficiency (Ward, 2015).

3.2.2 The gene variant *egl-1 (bc275)* leads to increased mRNA stability and a general block in apoptosis

The CRISPR/Cas9 generated *egl-1 (bc275)* variant is, at the DNA level, an insertion mutation within the *egl-1* 3'UTR, which adds an additional miR-58 family binding site and a mutated miR-35-42 family binding site. I was able to show that this variant leads to a complete block in apoptosis during embryogenesis. Utilizing single molecule RNA fluorescent in situ hybridization (smRNA FISH), I determined that *egl-1* mRNA is present in the *egl-1 (bc275)* mutant animals. Within the examined RID lineage, *egl-1 (bc275)* mRNA levels are increased in the RIDnb compared to wildtype. Also, within the daughter cell of the RIDnb which is destined to die, I was able to detect an increased level of mRNA. Here one could speculate that the stability of the mRNA is increased. It has been reported in mammalian cells, that 3'UTR length plays a role in mRNA stability, but the effect is opposite in differentiated and proliferating cells. In differentiated cells longer 3'UTRs lead to an increased stability (Ji *et al.*, 2009), while in proliferating cells shorter 3'UTRs with fewer microRNA targeting sites are preferred (Sandberg *et al.*, 2008), which can further lead to oncogene activation in cancer cells (Mayr and Bartel, 2009). This is further supported by the finding that in the daughter of the RIDnb which is programmed to die, I observed increased levels of *egl-1 (bc275)* mRNA copies compared to wildtype. It has been observed in miR-35 and miR-58 family mutant background, that *egl-1* mRNA copies within the RIDnb and the RID neuron have nearly same level as in wildtype (Sherrard *et al.*, 2017). This indicates that, in the dying daughter, regulation by microRNAs is not active anymore and the increased levels I observed might be due to an enhanced stability of the mRNA.

On the other hand, a different folding of the 3'UTR is predicted for *egl-1 (bc275)* variant, which may prevent decay of the mRNA by *trans*-acting elements, although their binding-sites for the miR-35 and miR-58 families are intact. In the RIDnb I observed increased levels of *egl-1 (bc275)* mRNA. It has been shown that in a miR-35 family and miR-58 family mutant background the *egl-1* mRNA levels in the RIDnb are even higher compared to what I have observed in the *egl-1 (bc275)* background (Sherrard *et al.*, 2017). This suggests that *trans*-acting factors, like microRNAs are capable of regulating the *egl-1 (bc275)* 3'UTR variant on mRNA level. This further supports the proposal that both the miR-35 and miR-58 family miRNAs keep the copy number of *egl-1* mRNAs below the threshold necessary to trigger death in the RIDnb (Sherrard *et al.*, 2017). The change in predicted folding in the *egl-1 (bc275)* 3'UTR variant might also play a role in translation initiation. For example, by impairing spatial formation of a “closed loop” structure necessary for translation initiation would cause a block in translation. This would explain the general block in apoptosis during embryogenesis but also in later stages as seen by counting the extra cells within the pharynx. Interestingly, when looking at the RID neuron, the *egl-1* mRNA levels of the variant *bc275* and wildtype are not significantly different, this has also been shown in miR-35 and miR-58 family mutant backgrounds (Sherrard *et al.*, 2017). While there is no evidence that *egl-1* mRNA is maintained in this cell, it indicates that *egl-1* expression is controlled by a process independent from the 3'UTR.

3.3 Fluorescent CIRSPR/Cas9 mediated tracing of distinct DNA sequences

During embryonic development, genomic organization is changing rapidly, for example genomic loci change their localization within the nucleus when transcriptionally active as compared to their transcriptionally silent state. Current methods for visualization of these localizations, such as FISH, allow sequence specific visualization of DNA and RNA in fixed samples. To overcome this the CRISPR/Cas9 system was recently modified to allow DNA and RNA visualization in living cells (B. Chen *et al.*, 2013; Anton *et al.*, 2014; Nelles *et al.*, 2016). I adapted the method for *C. elegans* using a codon-optimized version of the dCas9 fused to two GFP and nuclear targeting signals. Unfortunately, the expression of the generated *dCas9::gfp::gfp* fusion construct under control of the constitutively active *mai-2* promoter leads to embryonic lethality, even at low concentrations within the injection mix, the construct is highly expressed. It was therefore not possible to achieve a successful integration. Nevertheless, exchanging the promoter, in future experiments, to an inducible promoter, such as a heat shock promoter, would bypass ubiquitous expression and therefore lethality at the timepoint of interest. While there is toxicity associated with the generated construct in *C. elegans*, some cells expressed the fusion construct mosaically. In these cells, the fusion construct localizes to the nucleus and is enriched within the nucleoli. This indicates that the generated fusion construct is in general functional, this is further supported by the observation that the dCas9 can non-specifically interact with ribosomal RNA or nuclear proteins leading to nucleoli localization (Chen *et al.*, 2016). In cells which are expressing the construct, the *dCas9::gfp::gfp* fusion construct expression was too high, so that the telomere targeting sgRNAs became the limiting factor and the signals which might be present would be outshined. Nevertheless, integration of *dCas9::gfp::gfp* fusion construct into the genome is still advised, since it would allow one to easily switch the target, just by injection a different set of sgRNA.

3.3.1 Future applications

The method presented would allow deeper insights into the regulation of *egl-1*. First, the spatial localization of the *egl-1* locus would be an interesting starting point to look at, especially during embryonic development. One could speculate that the *egl-1* locus is transcriptionally silent prior to the first wave of cell death and is therefore closed and/or located outside of the transcriptionally active area within the nucleus. This could result in a low or absent signal due to the condensed state of the chromatin. Since *egl-1* is required for the first wave of cell death, the *egl-1* locus needs to be accessible at this timepoint. The *egl-1* locus could have been moved to a transcriptionally active area within the nucleus when compared to the closed locus position. Timeseries experiments of the embryonic development would give insights into the state before the first cell death, but it would be also interesting to examine the *egl-1* locus in fully differentiated cells and cells undergoing apoptosis. Here, the invariant lineage of *C. elegans* would allow for a dedicated view of the *egl-1* locus behavior in a single cell or a dedicated cell lineage.

Second, this method can also be adapted to be used for chromatin immunoprecipitation. Here either the GFP or the dCas9 could be directly targeted via an antibody to specifically precipitate the desired dCas9 target. This technique is applied in, the commercially available, engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP), which utilizes the CRISPR/Cas9 system. Here enChIP can be combined with mass spectrometry (enChIP-MS), next-generation sequencing (enChIP-Seq), and RNA-Seq (enChIP-RNA-Seq) (Fujita and Fujii, 2013; Fujita *et al.*, 2013). This would make it possible to identify proteins, interacting genomic regions, and RNA interaction partners, respectively.

For looking at protein-protein interactions at specific DNA-loci a look at the genomic locus proteomics (GLoPro) method might be interesting (Myers *et al.*, 2018). By fusion of dCas9 to ascorbate peroxidase (APEX2) followed by chromatin immunoprecipitation

and mass spectrometry, quantitative proteomics at a non-repetitive single locus has been shown in HEK293T cells. In *C. elegans* ascorbate peroxidase (APX) has been used for *in vivo* mapping of subcellular-tissues-specific proteomes (Reinke *et al.*, 2017), but a direct transfer of the GLoPro method has, to my knowledge, not been shown yet.

3.4 *In silico* identification of chromatin modifiers with potential impact on *egl-1*

3.4.1 Chromatin factor identification and quality control

One of the biggest weaknesses of any semantic approach is that it relies on maintenance and curation effort by the scientific community. Often semantic data, currently available, is lacking the latest state of scientific knowledge. In the case studied here, lookup of all chromatin modifiers was not possible, because a portion of the genes is not correctly assigned within the gene ontology or is missing relationships to their ontology terms. Further, scientific experiments in the past are often worked with gene or sequence names instead of identifiers of a gene ontology; here an intermediate mapping step is required. For the analysis I conducted, a fallback to text-based search was necessary, since it was not possible to get all genes of interests directly from the gene ontology. This has of course drawbacks, such as missing synonyms which could be resolved by an ontology; here I tried to compensate for this by generation of a word list including wildcard characters. This approach leads to false positive genes within the dataset, since keywords may also be part gene descriptions, in which for example an interaction with a chromatin identifier is described. This could be seen also in my dataset, where I identified apoptosis associated candidates which are not involved in chromatin modifications but do impact *egl-1*. Nevertheless, the keyword-based approach was able to identify the majority (63 of 68) of known *C. elegans* chromatin modifying genes. I have shown that the developed Python application is able to sort a given set of genes for best fitting to the idealized expression pattern of a chromatin modifier I defined. With little effort the ranking could be adapted for different characteristics with the expression data profile of genes. Based on the GO term enrichment analysis, which shows that apoptosis related genes are indeed enriched, I conclude that the ranking of genes I presented is a good starting point for future

experiments. The good news is that within the scientific community more and more metadata points are connected to the ontologies and a lot of effort is taken to also link experimental or even raw data directly. Meaning, by repeating this analysis in future the accuracy will increase.

3.4.2 Future experiments

The hypothesis that the *egl-1* locus is only opened shortly before the first wave of cell deaths and that this happens due to chromatin-modifying genes, is based on the observation that on the one hand no cell deaths can be observed beforehand in the embryo and on the increasing expression level of *egl-1*. A chromatin modifier that supports this process should also show a changing expression level shortly before the expression level change of *egl-1*. The candidates that were selected in the presented analysis all show a peak for this point in time (either up or down regulated), this offers a good starting point for the further downstream experiments.

Since changes in the cell death pattern can be detected microscopically, a good approach would be to start with RNAi experiments or knock-out mutants of the candidate genes. Here, the timepoint just before and after the first wave of cell deaths is particularly interesting and time-consuming screening of the entire embryonic development can be avoided. Candidates who show abnormalities in cell death could then be examined more closely by single cell lineaging to determine the influence on dedicated cells or cell lines. Additionally, this could be combined with the fluorescent CRISPR/Cas9 mediated tracing of the *egl-1* locus. Studies in *C. elegans* have shown that silent genes are enriched at the nuclear periphery (Chuang *et al.*, 2006; Meister *et al.*, 2010). Based on the described hypothesis the *egl-1* locus should be closed prior to the first wave cell deaths. When the first cell deaths arise the *egl-1* locus must be open in order to be expressed. This could be reflected by the position of the *egl-1* locus within the nucleus. Although, it was reported that during embryonic development positioning of silent genes seems to be little or absent (Meister *et al.*, 2010).

This phenotypic analysis could be supported by further studies on the *egl-1* locus at the molecular level. Here, the density of nucleosomes in the *egl-1* locus at the defined timepoints is particularly interesting. By comparison of the candidate gene with wildtype the density as well as the position of the nucleosomes could differ. Typically, isolated chromatin is cut using micrococcal nuclease (MNase), leaving DNA wrapped around a nucleosome intact. These can then be examined with various methods which offer different resolutions. For example, the indirect end labeling (for individual loci) or microarrays with low resolution, as well as high-resolution methods such as parallel sequencing. While the latter one offers a resolution of one bp it requires to sequence the whole genome if the *egl-1* locus is not isolated beforehand. Maybe to overcome this engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR to isolate the *egl-1* locus prior to MNase digestion and sequencing would be a solution, although to my knowledge this has not been shown yet.

4 Material and Methods

4.1 Strains and maintenance

The strains were grown on NGM plates seeded with *E. coli* strain OP50. Strains used for MosSCI integration were grown on *E. coli* strain HB101 seeded NGM plates. All strains were maintained at 20 °C. If strains were grown at other temperatures L4 larvae were shifted to the desired temperature and the progeny was used for the experiments. The maintenance of the strains is described in (Brenner, 1974).

Most of the strains can be purchased from the *Caenorhabditis Genetics Center* (CGC, University of Minnesota, Minneapolis, USA – the CGC is funded by NIH Office of Research Infrastructure Programs (P40 OD010440)).

The strain N2 (Bristol) was used as wildtype reference in this study.

Allele/	LG	Transgene	Chapter	Reference
<i>mIn1</i>	II	(balancer)	2.2	(Edgley <i>et al.</i> , 2001)
<i>nDf49</i>	II	<i>mir-42-44</i>	2.2	(Miska <i>et al.</i> , 2007)
<i>nDf50</i>	II	<i>mir-35-41</i>	2.2	(Miska <i>et al.</i> , 2007)
<i>ed-3</i>	III	<i>unc-119</i>	2.2	(Maduro and Pilgrim, 1995)
<i>nDf53</i>	III	<i>mir-80</i>	2.2	(Miska <i>et al.</i> , 2007) (Alvarez-Saavedra and Horvitz, 2010)
<i>n717</i>	IV	<i>ced-3</i>	2.2	(Shaham <i>et al.</i> , 1999)
<i>n3330</i>	V	<i>egl-1</i>	2.2	B. Conradt
<i>nDf54</i>	X	<i>mir-81-82</i>	2.2	(Miska <i>et al.</i> , 2007)
<i>n4640</i>	X	<i>mir-58.1</i>	2.2	(Miska <i>et al.</i> , 2007)
<i>cn64</i>	II	<i>dpy-10</i>	2.3	(Hosono, 1980)
<i>e2123</i>	III	<i>pha-1</i>	2.3	(Schnabel and Schnabel, 1990)
<i>bc275</i>	V	<i>egl-1</i>	2.3	This study
<i>gk119</i>	I	<i>emr-1</i>	2.4	(<i>C. elegans</i> Deletion Mutant Consortium, 2012)

Table 15. Alleles used in this work

Allele/	LG	Plasmid	Transgene	Reference
<i>bcSi25</i>	I	pBC1483	P _{mai-2gfp::h2b::mai-2} 3' UTR	This study (generated by M. Basch, N. Memar, and me)
<i>bcSi26</i>	I	pBC1484	P _{mai-2gfp::h2b::egl-1^{wt}} 3' UTR	This study (generated by M. Basch, N. Memar, and me)
<i>bcSi27</i>	I	pBC1485	P _{mai-2gfp::h2b::egl-1^{mut mir-35}} 3' UTR	This study (generated by M. Basch, N. Memar, and me)
<i>bcSi46</i>	I	pBC1653	P _{mai-2gfp::h2b::egl-1^{mut mir-58}} 3' UTR	This study
<i>bcSi47</i>	II	pBC1654	P _{mai-2gfp::h2b::egl-1^{mut mir-35 mir-58}} 3' UTR	This study
nEx1187	-	-	mir-35 (genomic) + sur-5::gfp	(Alvarez-Saavedra and Horvitz, 2010)
bqSi142	II	-	emr-1p::emr-1::mCherry	(Morales-Martinez, Dobrzynska and Askjaer, 2015)

Table 16. Transgenes, extrachromosomal arrays used in this work.

4.2 Cloning of the *egl-1* 3'UTR reporters with microRNA binding-site variants

The plasmids pBC1483 and pBC1484 were made by M. Basch, N. Memar, and designed by me, utilizing overlap-extension PCRs. First P_{mai-2} (1763 bp), defined as upstream sequence beginning at *mai-2* start codon until the transcription unit B0546.4 was amplified using Primers

5'-agatctGGAAAAATCGATA-3' (lowercase indicates BglII site) and

5'-GAAAAGTTCTTCTCCTTTACTCATTCTGAAAATTGAGTGAATTAG-3'.

The *gfp::h2b* fragment (1287 bp) was amplified from pCM1.35 (a gift from G. Seydoux; Addgene plasmid no. 17248) and subsequently fused to P_{mai-2} using Primers

5'-CTAATTCACCTCAATTTTCAGAATGAGTAAAGGAGAAGAACTTTT-3' and

5'-GCTCGCGTTCTTGTACTGCAAATTACTTGCTGGAAGTGTACTTG-3'.

For pBC1483 the *mai-2* 3'UTR (142bp) and for pBC1484 the *egl-1* 3'UTR was used. Amplification of *mai-2* 3'UTR was done using the Primers

5'-CAAGTACACTTCCAGCAAGTAATTTGCAGTACAAGAACGCGAGC-3' and

5'-TcttaagTCGCTAAAACTA-3' (lowercase indicates AflII site)

Amplification of *egl-1* 3'UTR was done using the primers

5'-CAAGTACACTTCCAGCAAGTAAGTGATCAAAATCTCCAACCTTTTC-3' and

5'-TcttaagAAAAAAACCATATTTATTATTAG-3' (lowercase indicates AflII site)

All fragments were ligated blunt ended into pBluescript II KS+ linearized with EcoRV. For pBC1485 the ligated *egl-1* 3'UTR was mutated utilizing PCR mediated site-directed mutagenesis, altering the miR-35 family binding site from 5'-

GATTTCATATAATACCCGGT-3' to CTAATCTCATACATCGCCGA-3'. Primers used for the PCR were

5'-TCTAATCTCATACATCGCCGAGTTTTTCTTCATTTGTGATTATTTTTC-3' and

5'-CTCGGCGATGTATGAGATTAGATGGTACAAATTGGAGAAAAG-3'.

From the intermediate pBluescript plasmids each fragment was excised using BglII/AflII double digestion and ligated into a BglII/AflII linearized pCFJ350 MosSCI vector (a gift from Erik Jorgensen; Addgene plasmid no. 34866). The final plasmids are based on pCFJ350 containing *P_{mai-2}h2b::gfp::mai-2* 3'UTR termed pBC1483, *P_{mai-2}h2b::gfp::egl-1* 3'UTR termed pBC1484 and *P_{mai-2}h2b::gfp::egl-1^{mut miR-35}* 3'UTR termed pBC1485.

Based on pBC1484 and pBC1485 I generated two plasmids in which the miR-58 family binding-site is also mutated pBC1653(*P_{mai-2}h2b::gfp::egl-1^{mut miR-58}* 3'UTR) and pBC1654 (*P_{mai-2}h2b::gfp::egl-1^{mut miR-35 miR-58}* 3'UTR), respectively.

Therefore, I mutated 5'-GAGAGATCGAAAA-3' to 5'-CACTGTACCATAT-3' using PCR mediated site-directed mutagenesis using Primers

5'-CACTGTACCATATATAATCACAAATGAAGAAAAACAC-3' and

5'-ATATGGTACAGTGCGTCTCCAACCTCCCCT-3'.

4.3 Single copy integration of *egl-1* 3'UTR reporter

For all five *egl-1* 3'UTR reporter transgenes I generated single-copy integrations into the *C. elegans* genome utilizing universal MosSCI (Mos1-mediated Single Copy Insertion) (Frøkjær-Jensen *et al.*, 2008, 2014) and germline microinjections (Mello *et al.*, 1991). For MosSCI integrations on LGI I utilized strain EG8078 [*oxTi185; unc-119(ed3)*] and for integrations on LG II strain EG6699 [*attTi5605; unc-119(ed3)*] Table 17. Both strains were grown on *E. coli* HB101 seeded NGM plates at 20 °C. Microinjection mixes are described in Table 18.

Linkage group	Transgene	Allele
LGI	pBC1483	<i>bcSi25</i>
	pBC1483	<i>bcSi26</i>
	pBC1485	<i>bcSi27</i>
	pBC1653	<i>bcSi46</i>
LGII	pBC1654	<i>bcSi47</i>

Table 17: Transgenes and Alleles generated

Plasmid	Description	Final concentration
pCFJ601	<i>Peft-3::transposase</i>	50 ng/μl
pMA122	<i>Phsp::peel-1</i>	10 ng/ul
pGH8	<i>Prab-3::mCherry</i> (Pan-neuronal)	10 ng/μl
pCFJ90	<i>Pmyo-2::mCherry</i> (pharynx muscle)	2.5 ng/μl
pCFJ104	<i>Pmyo-3::mCherry</i> (body muscle)	5 ng/μl
	Transgene in targeting vector	50 ng/μl

Table 18: Injection mix for MosSCI

After injection three injected worms were put on one fresh *E. coli* HB101 seeded NGM plate, the plates were kept at 25 °C until starvation. After starvation plates were screened for successfully integrated animals. Therefore, at least 10 L1 or dauer animals that move very well and do not have any of the fluorescent co-injection markers (mCherry fluorescence) were picked clonally on small NGM plates with OP50. The next generation was screened for homozygosity by looking for movement impaired animals. Homozygous plates were then screened for the presence of GFP fluorescence of the reporter construct. From GFP positive plates 10 animals were validated by PCR and sequencing.

Animals were lysed in 10 µl worm lysis buffer and then frozen by -80 °C for at least 10 minutes. Subsequently followed by incubation 65 °C for 1 hour and 95 °C for 20 minutes.

For validation by PCR a positive and negative control set of primers was used. For positive control an 1864 bp long fragment spanning from the genomic region into the transgene was selected using primers

5' – TTGTGCATCTATAATGTAGCACCACAATACAAG–3' and

5' – CATGGTCATAGCTGTTTCCTGCC –3'.

The negative control is spanning the integration site for the transgene resulting in a 1969 bp fragment if integration was not successful. For PCR amplification the primers

5' – ACCGGAAACCAAAGGACGAGAG–3' and

5' – ACGCCCAGGAGAACACGTTAG–3' were used.

Validation by sequencing was setup by using 1 µl of worm lysis and 1 µl of primer (see Table 19) and 5µl EB buffer (NEB; T1016L). PCR and Sequencing was done by an internal service provider.

Primer (5' – 3')
CGAATGAGACCCTTGTTAAAATC
CTGACGTCTTCTTCTGCTTCTTTTC
CCACTTCTTTCCGTTTTTCAGCC
GGAGACATCCCAATCCATCATTTTC
GAAGATGGAAGCGTTCAACTAG
CCATCTAATTCAACAAGAATTGGG
ATGAGTAAAGGAGAAGAAGTTTCACT
GGAAAAAATCGATAAAAAGTTGATTAGC
ACGCCCAGGAGAACACGTTAG

Table 19: Sequencing primers for 3'UTR reporter transgenes.

4.4 4D microscopy and lineage analysis of *C. elegans* embryonic development

L4 larvae were grown overnight at 25 °C, from young adults two- or four-cell embryos were extracted and mounted with 4.5 % agarose pads on a slide. The mounted embryos were protected with a cover slip. Water was added between slide and cover slip, the cover slip was then sealed with liquid Vaseline to prevent drying out of the embryos. 4D recordings were made using a Zeiss Axio Imager.M2 controlled by Time to Live software (Caenotec). A 25 images thick DIC z-stack was taken every 35 seconds at 25 °C throughout embryonic development as further described in (Schnabel *et al.*, 1997, 2006). Lineage analysis was done using Simi BioCell software (Simi Reality Motion Systems).

4.5 Single-molecule RNA fluorescent in situ hybridization

4.5.1 Probes

For the single-molecule RNA fluorescent in situ hybridization experiment (smRNA FISH) hybridization was performed by K. Ikegami as described in (Raj *et al.*, 2008; Raj and Tyagi, 2010). The Stellaris FISH probes (Biosearch Technologies) were designed

by R. Sherrard based on the mature mRNA of the target gene utilizing the Stellaris Probe Designer (version 4.2). The *egl-1* probes were TAMRA-labeled and used at 250 nM. The *unc-3::gfp* fusion probes were fluorescein-labeled and used at 500 nM.

Probes (5'to 3')	Probes (5'to 3')	Probes (5'to 3')
CGGTGTGAATGTTTTGGGTG	AAGAATCTTCACACGAGGAG	GTTGGAGATTTTGATCACTT
AAGAGAAAGTTAGAATACGAC	AAAAAATCCCGAGTCGTCGG	CATGGTACAAATTGGAGAA
AGCATCAGCATATCAACTGA	CGATGCTGCTGATCTCAGAG	CACCGGGTATTATGAGAA
ATCCGAAGAGGTTGAGGCAA	ATTGCTGCTAGCTTGGAGCC	TAATCACAAATGAAGAAAA
AAACGTTGGACATTGGTAGA	GAGCATCGAAGTCATCGC	TGGAGACGGAGAGATCGAAA
AAAACGGAAGATTGAACGTC	TGGGCCGAGTAGGACATCAT	TGGTACAAATATTGAGGGGA
ACATGTTCTTTTCGTTGT	GAAGAGGCTTCTGTCGGAAG	AATATGAGCAATAAAGGAC
GTCCTGAGACGAGGAGTA	GCGAAAAAGTCCAGAAGACG	

Table 20. Sequences of *egl-1* smRNA FISH probes

Probes (5'to 3')	Probes (5'to 3')	Probes (5'to 3')
CATTTTTCTGAGCTCGGTA	GTGTCCAAGAATGTTTCCAT	CGAGATGTCTGATGACAGCG
CTTACGCTTCTTCTTTGGAG	GTGAGTTATAGTTGTATTCC	GGACTTAGAAGTCAGAGGCA
TACTCATTTTTCTACCGGT	GTCTGCCATGATGTATACAT	AGAGCATGTAGGGATGTTGA
CCAGTGAAAAGTTCTTCTCC	CTTTGATTCCATTCTTTTGT	AAAATAGGGGGTGGGAGCAC
CCCATTAAACATCACCATCTA	CCATCTTCAATGTTGTGTCT	TTTCATTGTTAGAGGTGACT
CCTCTCCACTGACAGAAAAT	ATGGTCTGCTAGTTGAACGC	GGGAGGGAGCACAATTTTTT
GTAAGTTTTCCGTATGTTGC	CGCCAATTGGAGTATTTTGT	GTGTACACAGAACATTGTGT
GTAGTTTTCCAGTAGTGCAA	GTCTGGTAAAAGGACAGGGC	GTATTTTGTGTGCGGTTTTT
ACAAGTGTTGGCCATGGAAC	AAGGGCAGATTGTGTGGACA	GCGGTCATAAACTGAAACGT
GGTATCTCGAGAAGCATTGA	TCTTTTCGTTGGGATCTTTC	CCCAGACGTGCGAAGAAATA
TCATGCCGTTTCATATGATC	TCAAGAAGGACCATGTGGTC	CGATGAGCATGATTTGACGT
GGGCATGGCACTCTTGAAAA	AATCCCAGCAGCTGTTACAA	GGGGAACCCCAAAAAGCAA
TTCTTTCCTGTACATAACCT	TATAGTTCATCCATGCCATG	AAGAAAAACGCCGTCTCGA
GTTCCCGTCATCTTTGAAAA	GCTCAGTTGGAATTCTACGA	TGCATCGTGCTCATCAATAC
TGACTTCAGCACGTGTCTTG	CAAGTTGGTAATGGTAGCGA	CTCAAACCCAAACCTTCTTC
TAACAAGGGTATCACCTTCA	CCCTATTATTTTGGACACCA	ATCAACTTCTACTCACCTTC

Table 21. Sequences of *unc-3::gfp* smRNA FISH probes

4.5.2 Sample preparation

Embryos were harvested from adult worms by dissolving for ~8min in bleaching solution (NaHOCl 0.06 %, NaOH 0.7 N). Subsequently followed by pelleting embryos (1500 x g) and resuspension in M9 buffer (22 mM KH₂PO₄, 42 mM Na₂HPO₄, 86 mM NaCl, 1 mM MgSO₄). Embryos were then incubated at 25 °C until the desired stage. For fixation embryos were pelleted again (1500 x g) and resuspended in fixation solution (3.7 % Formaldehyde, 1x PBS) for 15 min at room temperature. For freeze-cracking the eggshell, embryos were submerged in liquid nitrogen for 1 min and thawed in water with room temperature, followed by 20 min incubation on ice. Embryos were washed twice with PBS and incubated overnight at 4 °C in 70 % ethanol. Embryos were resuspended in wash buffer (Formamide 10 %, SSC 2x) and incubated for 5 min at room temperature. For hybridization embryos were pelleted and resuspended in 100 µl hybridization buffer (Dextran sulfate 10 %, tRNA 1 mg / ml, Ribonucleoside Vanadyl Complex 2mM, BSA 200 µg / ml, SSC 2x Formamide 10 %) including the desired FISH probes. Embryos are then incubated overnight at 30 °C in the dark. On the next day embryos were washed twice in 1 ml wash buffer and incubated for 30 min at 30 °C. After incubation embryos are pelleted and incubated for 30 minutes at 30 °C in wash buffer with 5 ng / ml DAPI. Finally, embryos are pelleted and resuspended in VECTASHIELD Antifade Mounting Medium (Vector Laboratories; Cat. no. H-1000-10).

4.5.3 Imaging and quantification

Imaging was done using a LEICA TCS SP5 II confocal microscope with 63 x oil-immersion lens. The Leica LAS AF software was used with the following setting: pinhole 95.53 µm (Airy 1) line average 3, bidirectional scanning, 128 x 128 pixel (single cell), z-spacing 500 nm.

Images were analyzed using Fiji (Schindelin *et al.*, 2012), here a three dimensional region of interest (ROI) was defined for the target cell, then the background signal was

subtracted. Next, the smRNA FISH signal intensity was summed from z-projection summing all slices of the defined ROI. The mRNA copy number was calculated by division of the total signal intensity by the average intensity of a single diffraction-limited mRNA spot.

4.6 CRISPR/Cas9 and sgRNAs

4.6.1 Cloning sgRNAs

The sgRNAs targeting *pha-1*, *dpy-10*, *egl-1* (3'UTR) and telomeric repeats are based on the $P_{U6}::unc-119$ sgRNA (a gift from John Calarco; Addgene plasmid # 46169) (Friedland *et al.*, 2013). Both sgRNAs were generated site-directed mutagenesis with primers

5'-atgttcatcaagttattcatCAAACATTTAGATTTGCAATTC-3' and

5'-GatgaataacttgatgaacatGTTTTAGAGCTAGAAATAGCAAG-3' for *pha-1*

(lowercase letters indicate targeting sequence.) For the *dpy-10* sgRNA primers

5'-GctcgtggtgcctatggttagcGTTTTAGAGCTAGAAATAGCAAG-3' and

5'-gctaccataggcaccacgagCAAACATTTAGATTTGCAATTC-3' were used.

For *egl-1* the primers

5'-GggtacaaattggagaaaagtGTTTTAGAGCTAGAAATAGCAAG-3' and

5'-acttttctccaatttgtaccCAAACATTTAGATTTGCAATTC-3' were used.

For telomeric repeats the primers

5'-ggcttaggcttaggcttaggGTTTTAGAGCTAGAAATAGCAAG-3' and

5'-cctaagcctaagcctaagccAAACATTTAGATTGCAATTC-3' were used. For future experiments and to ensure stable sgRNA expression the sgRNA has been subsequently cloned into pCFJ350 resulting pBC1701.

4.6.2 Repair templates

The repair *egl-1* 3'UTR template was designed as ssDNA of 118 bp length ordered as HPLC purified oligo with sequence

5'-gagttggagacggagagatcgaaaaataatcacaatgaagaaaaaacaccgggtattatgagaaatcatggtacaaattggagaaaagttggagattttgatcacttaaaaagcgaa-3'.

The 80 bp dsDNA *pha-1* repair template was used from (Ward, 2015) with oligos sequences

5'-caaaatacgaatcgaagactcaaaaagagtatgctgtatgattacagatgttcatcaagttattcataaatcattgatag-3' and

5'-ctatcaatgatattatgaataacttgatgaacatctgtaatcatcacagcatactctttttgagtccttcgattcgtattttg-3'.

The repair template oligos were annealed by heating to 95 °C for 2 minutes and the slowly cooled down to 25 °C over 30 minutes. The 101 bp ssDNA *dpy-10* repair template was used from (Arribere *et al.*, 2014) with sequence

5'-CACTTGAACCTCAATACGGCAAGATGAGAATGACTGGAAACCGTACCGCATGCGGTGCCTATGGTAGCGGAGCTTCACATGGCTTCAGACCAACAGCCTAT-3'.

4.6.3 Cloning *dCas9::gfp::gfp* fusion construct

Based on pDD162 (was a gift from Bob Goldstein (Addgene plasmid # 47549) the RuvC1 (D10A) and HNH (H840A) endonuclease sites were mutated using site-directed mutagenesis with primers

5' -CAGCATCGGCCTTGCCATCGGAACGAACTC-3' and

5' -GAGTTCGTTCCGATGGCAAGGCCGATGCTG-3' for mutation of D10A.

5' -CTACGACGTCGACGCCATCGTCCCACAATC-3' and

5' -GATTGTGGGACGATGGCGTCGACGTCGTAG-3' for HNH (H840A) both mutations leading to inactivation form termed dCas9. The dCas9 was then subsequently amplified with overhanging primers for P_{mai-2}

5' - ctaattcactcaatttttcagaATGGACAAAAAATACAGCATCGG -3'

(lowercase letters indicate P_{mai-2} overhang) and

5' - gatccggatccGGCGTAGTCTGGGACGTCGT -3'

with overhang for the GS-Linker.

The GS-Linker and SV-40 NLS were ordered as oligo with overhang to the dCas9 and GFP. The oligos were annealed to dsDNA with sequence

5' -cagactacgccGGATCCGGATCCGGATCCGGATCCGGATCCCCAAAGAAGAAGCGTAAGG TCatgagtaaagg-3' (lowercase letters at 5' indicate dCas9 overhang and GFP overhang at 3').

The first GFP was amplified from pEZ151 (a gift from Esther Zanin) with overhangs to GS-Linker/SV40 NLS and the second GFP on the end using primers

5' - AGCGTAAGGTCatgagtaaaggagaagaact-3' and

5' -TTACTCATcggccctccgggcccctttgtatagttcatccatgc-3' .

The second GFP has different codon usage as compared to the first GFP to prevent recombination events. It was amplified from pJZ4 (a gift from Jeffrey Zielich) with overhangs to the first GFP and mai-2 3'UTR with primers

5'-ctagaccttacgcttcttcttttgCTGTAGAGCTCGTCCATTCCGTG-3' and

5'-aggccATGCTAGTggcccgaggggccgATG-3'.

Finally, I amplified the vector from pBC1483 (described above), with overhangs to the *dCas9* fragment and overhangs to the second *gfp* fragment with primers

5'-CGATGCTGTATTTTTGTCCATtctgaaaattgagtgaattagag-3' and

5'-CCAAAGAAGAAGCGTAAGGTCTAGtttgcagtacaagaacgcga-3'.

The vector and all four fragments have been ultimately assembled via Gibson cloning. Resulting $P_{mai-2}::dCas9::GS-Linker::SV40-NLS::gfp::gfp::SV40-NLS::mai-2$ 3'UTR listed as pBC1696.

4.6.4 Microinjection

All plasmids were purified using Qiagen's Mini Plasmid Purification kit. Injection mixes were injected into the gonad arms of young adult hermaphrodite worms as described before (Mello and Fire, 1995). For co-conversion experiments of *egl-1* and *pha-1*, *pha-1(e2123)* mutant worms were injected according to Table 22. For co-conversion of *egl-1* and *dpy-10* wildtype worms were injected according to Table 23.

Construct	Concentration
<i>egl-1</i> sgRNA	50 ng / μ l
<i>egl-1</i> repair template	5 ng / μ l
pDD162 ($P_{\text{eff-3}}$ Cas9 expression plasmid)	50 ng / μ l
<i>pha-1</i> sgRNA	50 ng / μ l
<i>pha-1</i> repair template	5 ng / μ l

Table 22: *pha-1 egl-1* co-Conversion injection mix

Construct	Concentration
<i>egl-1</i> sgRNA	50 ng / μ l
<i>egl-1</i> repair template	5 ng / μ l
pDD162 ($P_{\text{eff-3}}$ Cas9 expression plasmid)	50 ng / μ l
<i>dpy-10</i> sgRNA	50 ng / μ l
<i>dpy-10</i> repair template	5 ng / μ l

Table 23: *dpy-10 egl-1* co-Conversion injection mix

4.7 Bioinformatic methods and tools

4.7.1 APIs and requested data

The *C. elegans* gene ontology was queried in Python via the InterMine service, wormmine, from WormBase (version WS271) web site (<http://www.wormbase.org> and <http://intermine.wormbase.org/tools/wormmine/begin.do>). InterMine offers a pip package for API access to their data warehouse (<https://pypi.org/project/intermine/>).

Based on the keyword list (see Table 7) queries were executed against the gene ontology (using Taxon ID 6239 (*C. elegans*)). First, selecting of all GO term IDs associated with one or more keywords were executed. Second, all WormBase gene identifiers (WBGeneID) which are contained in the previously selected GO term IDs were requested and pooled. Third, WBGeneIDs were selected if the description or the brief description contain any of the keywords and added to the pool then all duplicates were removed. For quality control WBGeneIDs previously selected were compared to a list of known targets, in this case chromatin factors (see also chapter 2.5.3).

4.7.2 Expression data and analysis

The expression dataset was provided by Hashimshony et al. (Hashimshony *et al.*, 2015), which is spanning the whole embryonic development for all *C. elegans* genes.

The expression dataset was loaded into a data frame using pandas for loading excel sheets. Based on the before identified candidate list, expression data for candidate genes was extracted and the classification of candidate genes into the determined criteria was done as described below.

I defined a point of interest (POI) as variable (in this case expression data point 6) according to hypothesis described in chapter 2.5.4. Next I defined the classification scheme based on the idealized expression pattern I wanted to rank highest. Foreach

match I added a weight of one (point) to the ranking. By crossing a defined threshold extra weight was added as described.

I will here describe the procedure for genes with rising TPM before the POI and declining TPM after the POI. First, to determine if the transcription at the POI has the highest transcripts per million (TPM) in the time course, iteration over the individual gene dataset was done by comparing if each datapoint was smaller than the POI, then one point was added for the ranking. Second, the mean was calculated for the individual gene dataset, then it was determined if the average TPM is smaller than the POI TPM then one point was added to the ranking. Third, I compared if the TPMs at position POI-1 and POI+1 were at least 5 % smaller than at the POI and one point was added. If they are at least 10 % smaller two points were added and if they are 20 % smaller three points were added to the ranking. Fourth, or the TPMs of position POI-2 and POI+2 calculations were done as described in third. Fifth, it was determined if the TPM at position POI-1 was below the POI TPM and at position POI+1 above the POI TPM and one point was added. Sixth, for the TPMs of position POI-2 and POI+2 calculations were done as described in fifth. For genes with declining TPM before the POI and rising TPM after the POI the data is calculated conversely. The candidate genes are then ranked by the sum of points. It is important to note, that the 3rd or the 4th condition are not valid if the 5th or the 6th condition is true and vice versa. Further, the weight added for condition 5th and 6th should prefer raising or declining of TPMs over steady state TPMs.

All candidate genes which share the same ranking were exported in one sheet of an excel dataset, including their TPMs and ranking weights. Additionally, an overview sheet with all candidate genes was generated and enriched with metadata. Here the REST API of WormBase was queried (rest.wormbase.org) for each gene and the gene symbol and description was received.

4.7.3 Panther

The overrepresentation Test (Released 20190711) was done using PANTHER (Mi *et al.*, 2019) and the gene ontology ('Expansion of the Gene Ontology knowledgebase and resources', 2017). The parameters used were: Analysis Type: PANTHER Overrepresentation Test (Released 20190711); Annotation Version and Release Date: GO Ontology database (Released 2019-07-03); Reference List: *Caenorhabditis elegans* (all genes in database); Test Type: FISHER; Correction: FDR; GO biological process complete: *Caenorhabditis elegans* - REFLIST (19921).

5 References

- Ahringer, J. and Gasser, S. M. (2018) 'Repressive chromatin in *Caenorhabditis elegans*: Establishment, composition, and function', *Genetics*. Genetics Society of America, 208(2), pp. 491–511. doi: 10.1534/genetics.117.300386.
- Alvarez-Saavedra, E. and Horvitz, H. R. (2010) 'Many families of *C. elegans* microRNAs are not essential for development or viability.', *Current Biology*. Howard Hughes Medical Institute, 20(4), pp. 367–373. doi: 10.1016/j.cub.2009.12.051.
- Anders, C. *et al.* (2014) 'Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease', *Nature*, 513(7519), pp. 569–573. doi: 10.1038/nature13579.
- Andersen, E. C., Lu, X. and Horvitz, H. R. (2006) '*C. elegans* ISWI and NURF301 antagonize an Rb-like pathway in the determination of multiple cell fates', *Development*, 133(14), pp. 2695–2704. doi: 10.1242/dev.02444.
- Annis, R. P. *et al.* (2016) 'Mature neurons dynamically restrict apoptosis via redundant premitochondrial brakes.', *The FEBS journal*. NIH Public Access, 283(24), pp. 4569–4582. doi: 10.1111/febs.13944.
- Anton, T. *et al.* (2014) 'Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system.', *Nucleus (Austin, Tex.)*, 5(February 2015), pp. 163–72. doi: 10.4161/nucl.28488.
- Aravind, L. and Koonin, E. V (2001) 'Prokaryotic Homologs of the Eukaryotic DNA-End-Binding Protein Ku, Novel Domains in the Ku Protein and Prediction of a Prokaryotic Double-Strand Break Repair System', *Genome Research*, 11(8), pp. 1365–1374. doi: 10.1101/gr.181001.
- Arents, G. and Moudrianakis, E. N. (1995) 'The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 92(24), pp. 11170–4. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7479959> (Accessed: 2 September 2017).
- Arribere, J. A. *et al.* (2014) 'Efficient Marker-Free Recovery of Custom Genetic Modifications with CRISPR/Cas9 in *Caenorhabditis elegans*', *Genetics*, 198(3), pp. 837–846. doi: 10.1534/genetics.114.169730.

Bailis, A. M. and Rothstein, R. (1990) 'A defect in mismatch repair in *Saccharomyces cerevisiae* stimulates ectopic recombination between homeologous genes by an excision repair dependent process.', *Genetics*. Genetics Society of America, 126(3), pp. 535–47. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2249754> (Accessed: 31 August 2017).

Barrangou, R. *et al.* (2007) 'CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes', *Science*, 315(5819), pp. 1709–1712. doi: 10.1126/science.1138140.

Barres, B. A. and Raff, M. C. (1999) 'Axonal control of oligodendrocyte development.', *The Journal of cell biology*, 147(6), pp. 1123–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10601327> (Accessed: 18 August 2017).

Bartel, D. P. (2009) 'MicroRNAs: target recognition and regulatory functions.', *Cell*, 136(2), pp. 215–33. doi: 10.1016/j.cell.2009.01.002.

Benard, A. *et al.* (2014) 'DNA methylation of apoptosis genes in rectal cancer predicts patient survival and tumor recurrence.', *Apoptosis : an international journal on programmed cell death*. NIH Public Access, 19(11), pp. 1581–93. doi: 10.1007/s10495-014-1022-z.

Bernstein, E. *et al.* (2001) 'Role for a bidentate ribonuclease in the initiation step of RNA interference.', *Nature*, 409(6818), pp. 363–366. doi: 10.1038/35053110.

Bickmore, W. A. and Sumner, A. T. (1989) 'Mammalian chromosome banding — an expression of genome organization', *Trends in Genetics*, 5, pp. 144–148. doi: 10.1016/0168-9525(89)90055-3.

Bienroth, S., Keller, W. and Wahle, E. (1993) 'Assembly of a processive messenger RNA polyadenylation complex.', *The EMBO journal*. European Molecular Biology Organization, 12(2), pp. 585–94. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8440247> (Accessed: 9 February 2019).

Bilenoglu, O., Basak, A. N. and Russell, J. E. (2002) 'A 3'UTR mutation affects beta-globin expression without altering the stability of its fully processed mRNA', *British Journal of Haematology*, 119(4), pp. 1106–1114. doi: 10.1046/j.1365-2141.2002.03989.x.

Bohnsack, M. T., Czaplinski, K. and Gorlich, D. (2004) 'Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs.', *RNA (New York, N.Y.)*, 10(2), pp. 185–91. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/14730017> (Accessed: 30 August 2017).

Boise, L. H. *et al.* (1993) 'bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death.', *Cell*, 74(4), pp. 597–608. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8358789> (Accessed: 30 December 2018).

Breckenridge, D. G., Kang, B.-H. and Xue, D. (2009) 'Bcl-2 proteins EGL-1 and CED-9 do not regulate mitochondrial fission or fusion in *Caenorhabditis elegans*.', *Current biology : CB*. NIH Public Access, 19(9), pp. 768–73. doi: 10.1016/j.cub.2009.03.022.

Brenner, S. (1974) 'The genetics of *Caenorhabditis elegans*.', *Genetics*, 77(1), pp. 71–94. doi: 10.1002/cbic.200300625.

Broderick, J. A. *et al.* (2011) 'Argonaute protein identity and pairing geometry determine cooperativity in mammalian RNA silencing', *RNA*, 17(10), pp. 1858–1869. doi: 10.1261/rna.2778911.

Bucher, E. A. and Seydoux, G. (1994) 'Gastrulation in the nematode *Caenorhabditis elegans*', *Seminars in Developmental Biology*, 5(2), pp. 121–130. doi: 10.1006/sedb.1994.1016.

C. elegans Deletion Mutant Consortium (2012) 'Large-Scale Screening for Targeted Knockouts in the *Caenorhabditis elegans* Genome', *G3: Genes/Genomes/Genetics*, 2(11), pp. 1415–1425. doi: 10.1534/g3.112.003830.

CAI, X., Hagedorn, C. H. and Cullen, B. R. (2004) 'Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs', *RNA*, 10(12), pp. 1957–1966. doi: 10.1261/rna.7135204.

Carter, C. W. and Jr. (1978) 'Histone packing in the nucleosome core particle of chromatin.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 75(8), pp. 3649–53. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/278980> (Accessed: 2 September 2017).

Ceol, C. J. and Horvitz, H. . R. (2004) 'A New Class of *C. elegans* synMuv Genes Implicates a Tip60/NuA4-like HAT Complex as a Negative Regulator of Ras Signaling', *Developmental Cell*, 6(4), pp. 563–576. doi: 10.1016/S1534-5807(04)00065-6.

Chakraborty, S. *et al.* (2015) 'Engulfment pathways promote programmed cell death by enhancing the unequal segregation of apoptotic potential', *Nature Communications*, 6(1), p. 10126. doi: 10.1038/ncomms10126.

Chen, B. *et al.* (2013) 'Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system', *Cell*. Elsevier, 155(7), pp. 1479–1491. doi: 10.1016/j.cell.2013.12.001.

Chen, B. *et al.* (2016) 'Expanding the CRISPR imaging toolset with *Staphylococcus aureus* Cas9 for simultaneous imaging of multiple genomic loci', *Nucleic acids research*. Oxford University Press, 44(8), p. e75. doi: 10.1093/nar/gkv1533.

Chen, C., Fenk, L. A. and de Bono, M. (2013) 'Efficient genome editing in *Caenorhabditis elegans* by CRISPR-targeted homologous recombination.', *Nucleic acids research*, 41(20), p. e193. doi: 10.1093/nar/gkt805.

Chen, C. Y. and Shyu, A. B. (1995) 'AU-rich elements: characterization and importance in mRNA degradation.', *Trends in biochemical sciences*, 20(11), pp. 465–70. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8578590> (Accessed: 10 February 2019).

Chen, F. (2000) 'Translocation of *C. elegans* CED-4 to Nuclear Membranes During Programmed Cell Death', *Science*, 287(5457), pp. 1485–1489. doi: 10.1126/science.287.5457.1485.

Chen, F. *et al.* (2000) 'Translocation of *C. elegans* CED-4 to nuclear membranes during programmed cell death.', *Science (New York, N.Y.)*, 287(5457), pp. 1485–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10688797> (Accessed: 28 January 2019).

Chen, J.-M., Férec, C. and Cooper, D. N. (2006) 'A systematic analysis of disease-associated variants in the 3' regulatory regions of human protein-coding genes II: the importance of mRNA secondary structure in assessing the functionality of 3' UTR variants', *Human Genetics*, 120(3), pp. 301–333. doi: 10.1007/s00439-006-0218-x.

Chen, K. and Rajewsky, N. (2007) 'The evolution of gene regulation by transcription factors and microRNAs.', *Nature reviews. Genetics*, 8(2), pp. 93–103. doi: 10.1038/nrg1990.

Chen, Y.-Z. *et al.* (2013) 'Caspase-mediated activation of *Caenorhabditis elegans* CED-8 promotes apoptosis and phosphatidylserine externalization', *Nature Communications*, 4(1), p. 2726. doi: 10.1038/ncomms3726.

Chin-Sang, I. D. and Chisholm, A. D. (2000) 'Form of the worm: genetics of epidermal morphogenesis in *C. elegans*.', *Trends in genetics : TIG*. Elsevier, 16(12), pp.

544–51. doi: 10.1016/S0168-9525(00)02143-0.

Chinnaiyan, A. M. *et al.* (1997) 'Interaction of CED-4 with CED-3 and CED-9: a molecular framework for cell death.', *Science (New York, N.Y.)*, 275(5303), pp. 1122–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9027312> (Accessed: 28 January 2019).

Chiorazzi, M. *et al.* (2013) 'Related F-box proteins control cell death in *Caenorhabditis elegans* and human lymphoma', *Proceedings of the National Academy of Sciences*, 110(10), pp. 3943–3948. doi: 10.1073/pnas.1217271110.

Choy, E. Y.-W. *et al.* (2008) 'An Epstein-Barr virus-encoded microRNA targets PUMA to promote host cell survival.', *The Journal of experimental medicine*. The Rockefeller University Press, 205(11), pp. 2551–60. doi: 10.1084/jem.20072581.

Chuang, C. H. *et al.* (2006) 'Long-Range Directional Movement of an Interphase Chromosome Site', *Current Biology*. Curr Biol, 16(8), pp. 825–831. doi: 10.1016/j.cub.2006.03.059.

Coller, J. M., Gray, N. K. and Wickens, M. P. (1998) 'mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation.', *Genes & development*. Cold Spring Harbor Laboratory Press, 12(20), pp. 3226–35. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9784497> (Accessed: 10 February 2019).

Cong, L. Le *et al.* (2013) 'Multiplex genome engineering using CRISPR/Cas systems.', 339(6121). doi: 10.1126/science.1231143.

Conradt, B. (2009) 'Genetic control of programmed cell death during animal development', *Annu Rev Genet.*, 43(1), pp. 493–523. doi: 10.1146/annurev.genet.42.110807.091533.Genetic.

Conradt, B. and Horvitz, H. R. R. (1998) 'The *C. elegans* protein EGL-1 is required for programmed cell death and interacts with the Bcl-2-like protein CED-9.', *Cell*, 93(4), pp. 519–29. doi: 10.1016/S0092-8674(00)81182-4.

Conradt, B. and Horvitz, H. R. R. (1999) 'The TRA-1A sex determination protein of *C. elegans* regulates sexually dimorphic cell deaths by repressing the *egl-1* cell death activator gene.', *Cell*, 98(3), pp. 317–327. doi: 10.1016/S0092-8674(00)81961-3.

Conradt, B., Wu, Y.-C. and Xue, D. (2016) 'Programmed Cell Death During *Caenorhabditis elegans* Development', *Genetics*, 203(4), pp. 1533–1562. doi: 10.1534/genetics.115.186247.

Conradt, B. and Xue, D. (2005) 'Programmed cell death', in *WormBook*, ed. *The C. elegans Research*. doi: doi/10.1895/wormbook.

Craig, J. M. (2005) 'Heterochromatin? many flavours, common themes', *BioEssays*. Wiley Subscription Services, Inc., A Wiley Company, 27(1), pp. 17–28. doi: 10.1002/bies.20145.

Cui, M. and Han, M. (2007) *Roles of chromatin factors in C. elegans development*, *WormBook*, ed. *The C. elegans Research*. WormBook, ed. The C. elegans Research Community. doi: doi/10.1895/wormbook.1.139.1.

Datsenko, K. A. *et al.* (2012) 'Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system', *Nature Communications*, 3, p. 945. doi: 10.1038/ncomms1937.

Decker, C. J. and Parker, R. (1994) 'Mechanisms of mRNA degradation in eukaryotes.', *Trends in biochemical sciences*, 19(8), pp. 336–40. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7940679> (Accessed: 10 February 2019).

Deltcheva, E. *et al.* (2011) 'CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.', *Nature*, 471(7340), pp. 602–607. doi: 10.1038/nature09886.

Denli, A. M. *et al.* (2004) 'Processing of primary microRNAs by the Microprocessor complex', *Nature*, 432(7014), pp. 231–235. doi: 10.1038/nature03049.

Dever, T. E. (1999) 'Translation initiation: adept at adapting.', *Trends in biochemical sciences*, 24(10), pp. 398–403. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10500305> (Accessed: 30 December 2018).

Dickinson, D. J. *et al.* (2013) 'Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination.', *Nature methods*. Nature Publishing Group, 10(10), pp. 1028–34. doi: 10.1038/nmeth.2641.

Dickinson, D. J. and Goldstein, B. (2016) 'CRISPR-Based Methods for *Caenorhabditis elegans* Genome Engineering', *Genetics*. Genetics Society of America, 202(3), pp. 885–901. doi: 10.1534/genetics.115.182162.

Edgley, M. *et al.* (2001) 'LG II balancer chromosomes in *Caenorhabditis elegans*: mT1(II;III) and the mln1 set of dominantly and recessively marked inversions', *Molecular Genetics and Genomics*, 266(3), pp. 385–395. doi: 10.1007/s004380100523.

von Ehrenstein, G. and Schierenberg, E. (1980) *Nematodes as Biological Models*.

Volume 1. Edited by Z. B. M. Academic Press, New York.

Elkayam, E. *et al.* (2012) 'The Structure of Human Argonaute-2 in Complex with miR-20a', *Cell*, 150(1), pp. 100–110. doi: 10.1016/j.cell.2012.05.017.

Ellis, H. M. and Horvitz, H. R. (1986) 'Genetic control of programmed cell death in the nematode *C. elegans*.', *Cell*, 44(6), pp. 817–29. doi: 10.1016/0092-8674(86)90004-8.

Elmore, S. (2007) 'Apoptosis: a review of programmed cell death.', *Toxicologic pathology*. NIH Public Access, 35(4), pp. 495–516. doi: 10.1080/01926230701320337.

Equence, C. E. S. *et al.* (1998) 'Genome Sequence of the Nematode *C. elegans*: A Platform for Investigating Biology', *Science (New York, N.Y.)*, 282(5396), pp. 2012–2018. doi: 10.1126/science.282.5396.2012.

'Expansion of the Gene Ontology knowledgebase and resources' (2017) *Nucleic Acids Research*. Narnia, 45(D1), pp. D331–D338. doi: 10.1093/nar/gkw1108.

Fenton, K. (2015) 'The effect of cell death in the initiation of lupus nephritis.', *Clinical and experimental immunology*. Wiley-Blackwell, 179(1), pp. 11–6. doi: 10.1111/cei.12417.

Ferreira, H. C. *et al.* (2013) 'The shelterin protein POT-1 anchors *caenorhabditis elegans* telomeres through SUN-1 at the nuclear periphery', *Journal of Cell Biology*, 203(5), pp. 727–735. doi: 10.1083/jcb.201307181.

Finch, J. T. *et al.* (1977) 'Structure of nucleosome core particles of chromatin.', *Nature*, 269(5623), pp. 29–36. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/895884> (Accessed: 2 September 2017).

Fiori, M. E. *et al.* (2018) 'miR-663 sustains NSCLC by inhibiting mitochondrial outer membrane permeabilization (MOMP) through PUMA/BBC3 and BTG2', *Cell Death & Disease*, 9(2), p. 49. doi: 10.1038/s41419-017-0080-x.

Frank, C. A. *et al.* (2005) '*C. elegans* HAM-1 positions the cleavage plane and regulates apoptosis in asymmetric neuroblast divisions', *Developmental Biology*, 284(2), pp. 301–310. doi: 10.1016/j.ydbio.2005.05.026.

Friedland, A. E. *et al.* (2013) 'Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system.', *Nature methods*, 10(8), pp. 741–3. doi: 10.1038/nmeth.2532.

Friedman, R. C. *et al.* (2009) 'Most mammalian mRNAs are conserved targets of microRNAs.', *Genome research*, 19(1), pp. 92–105. doi: 10.1101/gr.082701.108.

Frøkjær-Jensen, C. *et al.* (2008) 'Single-copy insertion of transgenes in *Caenorhabditis elegans*.', *Nature Genetics*. NIH Public Access, 40(11). doi: 10.1038/ng.248.

Frøkjær-Jensen, C. *et al.* (2014) 'Random and targeted transgene insertion in *Caenorhabditis elegans* using a modified Mos1 transposon.', *Nature methods*. NIH Public Access, 11(5), pp. 529–34. doi: 10.1038/nmeth.2889.

Fuchs, Y. and Steller, H. (2011) 'Programmed cell death in animal development and disease.', *Cell*. NIH Public Access, 147(4), pp. 742–58. doi: 10.1016/j.cell.2011.10.033.

Fujita, T. *et al.* (2013) 'Identification of telomere-associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP)', *Scientific Reports*, 3(1), p. 3171. doi: 10.1038/srep03171.

Fujita, T. and Fujii, H. (2013) 'Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR', *Biochemical and Biophysical Research Communications*, 439(1), pp. 132–136. doi: 10.1016/j.bbrc.2013.08.013.

Fulda, S., Meyer, E. and Debatin, K.-M. (2002) 'Inhibition of TRAIL-induced apoptosis by Bcl-2 overexpression', *Oncogene*, 21(15), pp. 2283–2294. doi: 10.1038/sj.onc.1205258.

Furuichi, Y., LaFiandra, A. and Shatkin, A. J. (1977) '5'-Terminal structure and mRNA stability.', *Nature*, 266(5599), pp. 235–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/557727> (Accessed: 30 December 2018).

Galluzzi, L. *et al.* (2012) 'Molecular definitions of cell death subroutines: recommendations of the Nomenclature Committee on Cell Death 2012.', *Cell death and differentiation*. Nature Publishing Group, 19(1), pp. 107–20. doi: 10.1038/cdd.2011.96.

Granato, M., Schnabel, H. and Schnabel, R. (1994) 'pha-1, a selectable marker for gene transfer in *C.elegans*', *Nucleic Acids Research*, 22(9), pp. 1762–1763. doi: 10.1093/nar/22.9.1762.

Gregory, R. I. *et al.* (2004) 'The Microprocessor complex mediates the genesis of microRNAs', *Nature*, 432(7014), pp. 235–240. doi: 10.1038/nature03120.

Griffith, J. D. *et al.* (1999) 'Mammalian Telomeres End in a Large Duplex Loop', *Cell*, 97(4), pp. 503–514. doi: 10.1016/S0092-8674(00)80760-6.

Grimson, A. *et al.* (2007) 'MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing', *Molecular Cell*. NIH Public Access, 27(1), pp. 91–105. doi: 10.1016/j.molcel.2007.06.017.

Grishok, A. *et al.* (2001) 'Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing.', *Cell*, 106(1), pp. 23–34. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11461699> (Accessed: 30 August 2017).

Grosswendt, S. *et al.* (2014) 'Unambiguous Identification of miRNA:Target Site Interactions by Different Types of Ligation Reactions', *Molecular Cell*, 54(6), pp. 1042–1054. doi: 10.1016/j.molcel.2014.03.049.

Grote, P. and Conradt, B. (2006) 'The PLZF-like Protein TRA-4 Cooperates with the Gli-like Transcription Factor TRA-1 to Promote Female Development in *C. elegans*', *Developmental Cell*, 11(4), pp. 561–573. doi: 10.1016/j.devcel.2006.07.015.

Gruber, A. R. *et al.* (2008) 'The Vienna RNA Websuite', *Nucleic Acids Research*, 36(Web Server), pp. W70–W74. doi: 10.1093/nar/gkn188.

Gu, J. and Lieber, M. R. (2008) 'Mechanistic flexibility as a conserved theme across 3 billion years of nonhomologous DNA end-joining', *Genes & Development*, 22(4), pp. 411–415. doi: 10.1101/gad.1646608.

Guhaniyogi, J. and Brewer, G. (2001) 'Regulation of mRNA stability in mammalian cells.', *Gene*, 265(1–2), pp. 11–23. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11255003> (Accessed: 10 February 2019).

Hale, M. L., Thapa, I. and Ghera, D. (2019) 'FunSet: an open-source software and web server for performing and displaying Gene Ontology enrichment analysis.', *BMC bioinformatics*. BioMed Central, 20(1), p. 359. doi: 10.1186/s12859-019-2960-9.

Hammond, S. M. *et al.* (2001) 'Argonaute2, a Link Between Genetic and Biochemical Analyses of RNAi', *Science*, 293(5532), pp. 1146–1150. doi: 10.1126/science.1064023.

Han, J. *et al.* (2004) 'The Drosha-DGCR8 complex in primary microRNA processing', *Genes & Development*, 18(24), pp. 3016–3027. doi: 10.1101/gad.1262504.

Harrison, M. M. *et al.* (2006) 'Some *C. elegans* class B synthetic multivulva proteins

encode a conserved LIN-35 Rb-containing complex distinct from a NuRD-like complex', *Proceedings of the National Academy of Sciences*, 103(45), pp. 16782–16787. doi: 10.1073/pnas.0608461103.

Harrow, J. *et al.* (2006) 'GENCODE: producing a reference annotation for ENCODE.', *Genome Biology*, 7(Suppl 1), p. S4. doi: 10.1186/gb-2006-7-s1-s4.

Hashimshony, T. *et al.* (2015) 'Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer', *Nature*. Nature Publishing Group, 519(7542), pp. 219–222. doi: 10.1038/nature13996.

Hatzold, J. and Conradt, B. (2008) 'Control of apoptosis by asymmetric cell division.', *PLoS Biology*. Edited by J. Ahringer. Public Library of Science, 6(4), p. e84. doi: 10.1371/journal.pbio.0060084.

Heitz, E. (1928) 'Das Heterochromatin der Moose', *Jahrbücher für wissenschaftliche Botanik*, pp. 762–818.

Heler, R. *et al.* (2015) 'Cas9 specifies functional viral targets during CRISPR-Cas adaptation.', *Nature*, 519(7542), pp. 199–202. doi: 10.1038/nature14245.

Hengartner, M. O. and Horvitz, H. R. R. (1994) 'C. elegans cell survival gene ced-9 encodes a functional homolog of the mammalian proto-oncogene bcl-2', *Cell*, 76(4), pp. 665–76. doi: 10.1016/0092-8674(94)90506-1.

Hengartner, M. O. M. O. *et al.* (1992) 'Caenorhabditis elegans gene ced-9 protects cells from programmed cell death.', *Nature*, 356(6369), pp. 133–135. doi: 10.1038/356494a0.

Hinz, J., Laughery, M. and Wyrick, J. (2015) 'Nucleosomes inhibit Cas9 endonuclease activity in vitro', *Biochemistry*, 54, pp. 7063–7066. doi: 10.1021/acs.biochem.5b01108.

Hirose, T., Galvin, B. D. and Horvitz, H. R. (2010) 'Six and Eya promote apoptosis through direct transcriptional activation of the proapoptotic BH3-only gene egl-1 in Caenorhabditis elegans.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 107(35), pp. 15479–84. doi: 10.1073/pnas.1010023107.

Hirose, T. and Horvitz, H. R. (2013) 'An Sp1 transcription factor coordinates caspase-dependent and -independent apoptotic pathways.', *Nature*. Howard Hughes Medical Institute, 500(7462). doi: 10.1038/nature12329.

Hoeppner, D. J., Hengartner, M. O. and Schnabel, R. (2001) 'Engulfment genes cooperate with ced-3 to promote cell death in *Caenorhabditis elegans*', *Nature*, 412(6843), pp. 202–206. doi: 10.1038/35084103.

Horlbeck, M. a *et al.* (2016) 'Nucleosomes impede Cas9 access to DNA in vivo and in vitro', *eLife*, 5, p. e12677. doi: 10.7554/eLife.12677.

Hosono, R. (1980) 'A study of morphology of *Caenorhabditis elegans*: A mutant of *Caenorhabditis elegans* with dumpy and temperature-sensitive roller phenotype', *Journal of Experimental Zoology*. John Wiley & Sons, Ltd, 213(1), pp. 61–67. doi: 10.1002/jez.1402130109.

Huang, W. *et al.* (2013) 'Mechanistic insights into CED-4-mediated activation of CED-3', 4, pp. 2039–2048. doi: 10.1101/gad.224428.113.Freely.

Hughes, T. A. (2006) 'Regulation of gene expression by alternative untranslated regions', *Trends in Genetics*, 22(3), pp. 119–122. doi: 10.1016/j.tig.2006.01.001.

Hutvagner, G. *et al.* (2001) 'A Cellular Function for the RNA-Interference Enzyme Dicer in the Maturation of the let-7 Small Temporal RNA', *Science*, 293(5531), pp. 834–838. doi: 10.1126/science.1062961.

Irmeler, M. *et al.* (1997) 'Direct physical interaction between the *Caenorhabditis elegans* "death proteins" CED-3 and CED-4.', *FEBS letters*, 406(1–2), pp. 189–90. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9109415> (Accessed: 30 January 2019).

Ishino, Y. *et al.* (1987) 'Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product.' American Society for Microbiology (ASM), 169(12). doi: 10.1128/jb.169.12.5429-5433.1987.

Isik, M., Korswagen, H. C. and Berezikov, E. (2010) 'Expression patterns of intronic microRNAs in *Caenorhabditis elegans*.' *Silence*. BioMed Central, 1(1), p. 5. doi: 10.1186/1758-907X-1-5.

James, C. *et al.* (1997) 'CED-4 induces chromatin condensation in *Schizosaccharomyces pombe* and is inhibited by direct physical association with CED-9', *Current Biology*, 7(4), pp. 246–252. doi: 10.1016/S0960-9822(06)00120-5.

Jan, C. H. *et al.* (2011) 'Formation, regulation and evolution of *Caenorhabditis elegans* 3'UTRs.', *Nature*. NIH Public Access, 469(7328), pp. 97–101. doi:

10.1038/nature09616.

Jansen, R. *et al.* (2002) 'Identification of genes that are associated with DNA repeats in prokaryotes.', *Molecular microbiology*, 43(6), pp. 1565–75. doi: 10.1046/j.1365-2958.2002.02839.x.

Jeffares, D. C., Penkett, C. J. and Bähler, J. (2008) 'Rapidly regulated genes are intron poor', *Trends in Genetics*, 24(8), pp. 375–378. doi: 10.1016/j.tig.2008.05.006.

Ji, Z. *et al.* (2009) 'Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development', *Proceedings of the National Academy of Sciences*, 106(17), pp. 7028–7033. doi: 10.1073/pnas.0900028106.

Jinek, M. *et al.* (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.', *Science (New York, N.Y.)*, 337(6096), pp. 816–21. doi: 10.1126/science.1225829.

Johnson, S. M. *et al.* (2005) 'RAS Is Regulated by the let-7 MicroRNA Family', *Cell*, 120(5), pp. 635–647. doi: 10.1016/j.cell.2005.01.014.

Jorgensen, I. and Miao, E. A. (2015) 'Pyroptotic cell death defends against intracellular pathogens.', *Immunological reviews*, 265(1), pp. 130–42. doi: 10.1111/imr.12287.

Jorgensen, I., Rayamajhi, M. and Miao, E. A. (2017) 'Programmed cell death as a defence against infection', *Nature Reviews Immunology*, 17(3), pp. 151–164. doi: 10.1038/nri.2016.147.

Kagias, K. and Pocock, R. (2015) 'microRNA regulation of the embryonic hypoxic response in *Caenorhabditis elegans*.', *Scientific reports*. Nature Publishing Group, 5, p. 11284. doi: 10.1038/srep11284.

Kapranov, P. *et al.* (2002) 'Large-scale transcriptional activity in chromosomes 21 and 22.', *Science (New York, N.Y.)*, 296(5569), pp. 916–9. doi: 10.1126/science.1068597.

Karvelis, T., Gasiunas, G. and Siksnys, V. (2017) 'Methods for decoding Cas9 protospacer adjacent motif (PAM) sequences: A brief overview', *Methods*, 121–122, pp. 3–8. doi: 10.1016/j.ymeth.2017.03.006.

Kato, M. *et al.* (2009) 'Dynamic expression of small non-coding RNAs, including

novel microRNAs and piRNAs/21U-RNAs, during *Caenorhabditis elegans* development.’, *Genome biology*. BioMed Central, 10(5), p. R54. doi: 10.1186/gb-2009-10-5-r54.

Kawamata, T. and Tomari, Y. (2010) ‘Making RISC’, *Trends in Biochemical Sciences*, 35(7), pp. 368–376. doi: 10.1016/j.tibs.2010.03.009.

Kerr, J. F., Wyllie, A. H. and Currie, A. R. (1972) ‘Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics.’ Nature Publishing Group, 26(4). doi: 10.1038/bjc.1972.33.

Kiontke, K. and Sudhaus, W. (2006) ‘Ecology of *Caenorhabditis* species’, *WormBook*, ed. *The C. elegans Research Community*. Available at: http://www.wormbook.org/chapters/www_ecolCaenorhabditis/ecolCaenorhabditis.html.

Kroemer, G. *et al.* (2009) ‘Classification of cell death: recommendations of the Nomenclature Committee on Cell Death 2009’, *Cell Death & Differentiation*, 16(1), pp. 3–11. doi: 10.1038/cdd.2008.150.

Kusenda, B. *et al.* (2006) ‘MicroRNA biogenesis, functionality and cancer relevance.’, *Biomedical papers of the Medical Faculty of the University Palack??, Olomouc, Czechoslovakia*, 150(2), pp. 205–15. doi: 10.5507/bp.2006.029.

Labi, V. and Erlacher, M. (2015) ‘How cell death shapes cancer.’, *Cell death & disease*. Nature Publishing Group, 6(3), p. e1675. doi: 10.1038/cddis.2015.20.

Lall, S., Piano, F. and Davis, R. E. (2005) ‘*Caenorhabditis elegans* decapping proteins: localization and functional analysis of Dcp1, Dcp2, and DcpS during embryogenesis.’, *Molecular biology of the cell*. American Society for Cell Biology, 16(12), pp. 5880–90. doi: 10.1091/mbc.e05-07-0622.

Lambie, E. J. and Conradt, B. (2016) ‘Deadly dowry: how engulfment pathways promote cell killing’, *Cell Death & Differentiation*, 23(4), pp. 553–554. doi: 10.1038/cdd.2015.170.

Large, E. E. and Mathies, L. D. (2007) ‘Chromatin regulation and sex determination in *Caenorhabditis elegans*’, *Trends in Genetics*, 23(7), pp. 314–317. doi: 10.1016/j.tig.2007.04.002.

Lau, N. C. (2001) ‘An Abundant Class of Tiny RNAs with Probable Regulatory Roles in *Caenorhabditis elegans*’, *Science*, 294(5543), pp. 858–862. doi: 10.1126/science.1065062.

Le, H. D. *et al.* (2004) 'A high proportion of genes involved in position effect variegation also affect chromosome inheritance', *Chromosoma*, 112(6), pp. 269–276. doi: 10.1007/s00412-003-0272-2.

Lee, R. C., Feinbaum, R. L. and Ambros, V. (1993) 'The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*.', *Cell*, 75(5), pp. 843–54. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8252621> (Accessed: 10 November 2013).

Lee, Y. *et al.* (2003) 'The nuclear RNase III Drosha initiates microRNA processing', *Nature*, 425(6956), pp. 415–419. doi: 10.1038/nature01957.

Lee, Y. *et al.* (2004) 'MicroRNA genes are transcribed by RNA polymerase II', *The EMBO Journal*, 23(20), pp. 4051–4060. doi: 10.1038/sj.emboj.7600385.

Lettre, G. *et al.* (2004) 'Genome-wide RNAi identifies p53-dependent and -independent regulators of germ cell apoptosis in *C. elegans*', *Cell Death and Differentiation*, 11(11), pp. 1198–1203. doi: 10.1038/sj.cdd.4401488.

Lewis, B. P., Burge, C. B. and Bartel, D. P. (2005) 'Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.', *Cell*, 120(1), pp. 15–20. doi: 10.1016/j.cell.2004.12.035.

Lewis, B. P., Green, R. E. and Brenner, S. E. (2003) 'Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans', *Proceedings of the National Academy of Sciences*, 100(1), pp. 189–192. doi: 10.1073/pnas.0136770100.

Lewis, J. D., Gunderson, S. I. and Mattaj, I. W. (1995) 'The influence of 5' and 3' end structures on pre-mRNA metabolism.', *Journal of cell science. Supplement*, 19, pp. 13–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8655642> (Accessed: 10 February 2019).

Lewis, J. D. and Izaurilde, E. (1997) 'The Role of the Cap Structure in RNA Processing and Nuclear Export', *European Journal of Biochemistry*. John Wiley & Sons, Ltd (10.1111), 247(2), pp. 461–469. doi: 10.1111/j.1432-1033.1997.00461.x.

Li, H. J. (1975) 'A model for chromatin structure.', *Nucleic acids research*, 2(8), pp. 1275–89. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1101222> (Accessed: 2 September 2017).

Lieber, M. R. (2010) 'The mechanism of double-strand DNA break repair by the

nonhomologous DNA end-joining pathway.', *Annual review of biochemistry*. NIH Public Access, 79, pp. 181–211. doi: 10.1146/annurev.biochem.052308.093131.

Lindsten, T. *et al.* (2000) 'The combined functions of proapoptotic Bcl-2 family members bak and bax are essential for normal development of multiple tissues.', *Molecular cell*, 6(6), pp. 1389–99. doi: 10.1016/S1097-2765(00)00136-2.

Liu, H. *et al.* (2006) 'Direct regulation of egl-1 and of programmed cell death by the Hox protein MAB-5 and by CEH-20, a *C. elegans* homolog of Pbx1', *Development*, 133(4), pp. 641–650. doi: 10.1242/dev.02234.

Liu, Q. A. and Hengartner, M. O. (1998) 'Candidate adaptor protein CED-6 promotes the engulfment of apoptotic cells in *C. elegans*.', *Cell*, 93(6), pp. 961–72. doi: 10.1016/S0092-8674(00)81202-7.

Lorenz, R. *et al.* (2011) 'ViennaRNA Package 2.0', *Algorithms for Molecular Biology*, 6(1), p. 26. doi: 10.1186/1748-7188-6-26.

Lu, J. *et al.* (2005) 'MicroRNA expression profiles classify human cancers', *Nature*, 435(7043), pp. 834–838. doi: 10.1038/nature03702.

Lu, X. and Horvitz, H. R. (1998) 'lin-35 and lin-53, Two Genes that Antagonize a *C. elegans* Ras Pathway, Encode Proteins Similar to Rb and Its Binding Protein RbAp48', *Cell*, 95(7), pp. 981–991. doi: 10.1016/S0092-8674(00)81722-5.

Luger, K. *et al.* (1997) 'Crystal structure of the nucleosome core particle at 2.8 Å resolution', *Nature*. Nature Publishing Group, 389(6648), pp. 251–260. doi: 10.1038/38444.

Lund, E. *et al.* (2004) 'Nuclear Export of MicroRNA Precursors', *Science*, 303(5654), pp. 95–98. doi: 10.1126/science.1090599.

MacRae, I. J. *et al.* (2006) 'Structural Basis for Double-Stranded RNA Processing by Dicer', *Science*, 311(5758), pp. 195–198. doi: 10.1126/science.1121638.

Maduro, M. and Pilgrim, D. (1995) 'Identification and cloning of unc-119, a gene expressed in the *Caenorhabditis elegans* nervous system', *Genetics*, 141(3), pp. 977–988.

Mali, P., Esvelt, K. M. and Church, G. M. (2013) 'Cas9 as a versatile tool for engineering biology.', *Nature methods*, 10(10), pp. 957–63. doi: 10.1038/nmeth.2649.

Martin, L. J. (2010) 'Mitochondrial and Cell Death Mechanisms in Neurodegenerative Diseases.', *Pharmaceuticals (Basel, Switzerland)*. Multidisciplinary Digital Publishing Institute (MDPI), 3(4), pp. 839–915. doi: 10.3390/ph3040839.

Martin, S. J. and Cotter, T. G. (1991) 'Ultraviolet B irradiation of human leukaemia HL-60 cells in vitro induces apoptosis.', *International journal of radiation biology*, 59(4), pp. 1001–16. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1674267> (Accessed: 18 August 2017).

Martinez, N. J. *et al.* (2008) 'Genome-scale spatiotemporal analysis of *Caenorhabditis elegans* microRNA promoter activity', *Genome Research*, 18(12), pp. 2005–2015. doi: 10.1101/gr.083055.108.

Masamha, C. P. *et al.* (2014) 'CFIm25 links alternative polyadenylation to glioblastoma tumour suppression.', *Nature*. NIH Public Access, 510(7505), pp. 412–6. doi: 10.1038/nature13261.

Massirer, K. B. *et al.* (2012) 'The miR-35-41 Family of MicroRNAs Regulates RNAi Sensitivity in *Caenorhabditis elegans*', *PLoS Genetics*. Edited by M. T. McManus. Public Library of Science, 8(3), p. e1002536. doi: 10.1371/journal.pgen.1002536.

Mayr, C. and Bartel, D. P. (2009) 'Widespread Shortening of 3'UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells', *Cell*, 138(4), pp. 673–684. doi: 10.1016/j.cell.2009.06.016.

McClintock, B. (1938) 'The fusion of broken ends of sister half-chromatids following chromatid breakage at meiotic anaphases'. Available at: <http://repository.cshl.edu/34583/> (Accessed: 2 September 2017).

McGlinchy, N. J. and Smith, C. W. J. (2008) 'Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense?', *Trends in Biochemical Sciences*, 33(8), pp. 385–393. doi: 10.1016/j.tibs.2008.06.001.

McJunkin, K. and Ambros, V. (2014) 'The embryonic mir-35 family of microRNAs promotes multiple aspects of fecundity in *Caenorhabditis elegans*.', *G3 (Bethesda, Md.)*. Genetics Society of America, 4(9), pp. 1747–54. doi: 10.1534/g3.114.011973.

Meisner, N.-C. *et al.* (2004) 'mRNA Openers and Closers: Modulating AU-Rich Element-Controlled mRNA Stability by a Molecular Switch in mRNA Secondary Structure', *ChemBioChem*, 5(10), pp. 1432–1447. doi: 10.1002/cbic.200400219.

Meister, P. *et al.* (2010) 'The spatial dynamics of tissue-specific promoters during C.

elegans development', *Genes and Development*. Cold Spring Harbor Laboratory Press, 24(8), pp. 766–782. doi: 10.1101/gad.559610.

Mello, C. C. *et al.* (1991) 'Efficient gene transfer in *C.elegans*: extrachromosomal maintenance and integration of transforming sequences.', *The EMBO journal*. European Molecular Biology Organization, 10(12), pp. 3959–70. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1935914> (Accessed: 19 October 2019).

Mello, C. and Fire, A. (1995) 'DNA transformation.', *Methods in cell biology*, 48, pp. 451–82. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8531738> (Accessed: 22 October 2019).

Merrick, W. C. (2004) 'Cap-dependent and cap-independent translation in eukaryotic systems', *Gene*, 332, pp. 1–11. doi: 10.1016/j.gene.2004.02.051.

Meyer, S., Temme, C. and Wahle, E. (2004) 'Messenger RNA Turnover in Eukaryotes: Pathways and Enzymes', *Critical Reviews in Biochemistry and Molecular Biology*, 39(4), pp. 197–216. doi: 10.1080/10409230490513991.

Meyne, J., Ratliff, R. L. and Moyzis, R. K. (1989) 'Conservation of the human telomere sequence (TTAGGG)_n among vertebrates.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 86(18), pp. 7049–53. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2780561> (Accessed: 2 September 2017).

Mi, H. *et al.* (2019) 'Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0)', *Nature Protocols*. Nature Publishing Group, 14(3), pp. 703–721. doi: 10.1038/s41596-019-0128-8.

Minn, A. J. *et al.* (1995) 'Expression of bcl-xL can confer a multidrug resistance phenotype.', *Blood*, 86(5), pp. 1903–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7655019> (Accessed: 14 August 2017).

Mishra, N., Wei, H. and Conradt, B. (2018) 'Caenorhabditis elegans ced-3 Caspase Is Required for Asymmetric Divisions That Generate Cells Programmed To Die.', *Genetics*. Genetics Society of America, 210(3), pp. 983–998. doi: 10.1534/genetics.118.301500.

Miska, E. A. *et al.* (2007) 'Most *Caenorhabditis elegans* microRNAs are individually not essential for development or viability', *PLoS Genetics*, 3(12), pp. 2395–2403. doi: 10.1371/journal.pgen.0030215.

Morales-Martinez, A., Dobrzynska, A. and Askjaer, P. (2015) 'Inner nuclear membrane protein LEM-2 is required for correct nuclear separation and morphology in *C. elegans*', *Journal of Cell Science*, 128(6), pp. 1090–1096. doi: 10.1242/jcs.164202.

Moteki, S. and Price, D. (2002) 'Functional coupling of capping and transcription of mRNA.', *Molecular cell*, 10(3), pp. 599–609. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12408827> (Accessed: 30 December 2018).

Muller (1938) 'The remaking of chromosomes', *Collecting Net*, 13. Available at: <http://www.citeulike.org/user/miketaboski/article/5837777> (Accessed: 2 September 2017).

Myers, S. A. *et al.* (2018) 'Discovery of proteins associated with a predefined genomic locus via dCas9-APEX-mediated proximity labeling', *Nature Methods*. Nature Publishing Group, 15(6), pp. 437–439. doi: 10.1038/s41592-018-0007-1.

Nakagawa, A., Sullivan, K. D. and Xue, D. (2014) 'Caspase-activated phosphoinositide binding by CNT-1 promotes apoptosis by inhibiting the AKT pathway', *Nature Structural & Molecular Biology*, 21(12), pp. 1082–1090. doi: 10.1038/nsmb.2915.

Nelles, D. A. *et al.* (2016) 'Programmable RNA Tracking in Live Cells with CRISPR/Cas9.', *Cell*, 165(2), pp. 488–96. doi: 10.1016/j.cell.2016.02.054.

Nelson, W. G. and Kastan, M. B. (1994) 'DNA strand breaks: the DNA template alterations that trigger p53-dependent DNA damage response pathways.', *Molecular and cellular biology*, 14(3), pp. 1815–23. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8114714> (Accessed: 18 August 2017).

Nikitina, T. and Woodcock, C. L. (2004) 'Closed chromatin loops at the ends of chromosomes.', *The Journal of cell biology*. The Rockefeller University Press, 166(2), pp. 161–5. doi: 10.1083/jcb.200403118.

Nishimasu, H. *et al.* (2014) 'Crystal structure of Cas9 in complex with guide RNA and target DNA.', *Cell*, 156(5), pp. 935–49. doi: 10.1016/j.cell.2014.02.001.

Nuñez, J. K. *et al.* (2015) 'Integrase-mediated spacer acquisition during CRISPR–Cas adaptive immunity', *Nature*, 519(7542), pp. 193–198. doi: 10.1038/nature14237.

O'Neil, N. and Rose, A. (2006) 'DNA Repair', *WormBook*, pp. 1–12. doi: 10.1895/wormbook.1.54.1.

Okamura, K. *et al.* (2004) 'Distinct roles for Argonaute proteins in small RNA-directed

RNA cleavage pathways', *Genes & Development*, 18(14), pp. 1655–1666. doi: 10.1101/gad.1210204.

Olins, A. L. and Olins, D. E. (1974) 'Spheroid chromatin units (v bodies).', *Science (New York, N.Y.)*, 183(4122), pp. 330–2. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/4128918> (Accessed: 2 September 2017).

Olsen, P. H. and Ambros, V. (1999) 'The lin-4 Regulatory RNA Controls Developmental Timing in *Caenorhabditis elegans* by Blocking LIN-14 Protein Synthesis after the Initiation of Translation', *Developmental Biology*, 216(2), pp. 671–680. doi: 10.1006/dbio.1999.9523.

van der Oost, J. *et al.* (2014) 'Unravelling the structural and mechanistic basis of CRISPR–Cas systems', *Nature Reviews Microbiology*, 12(7), pp. 479–492. doi: 10.1038/nrmicro3279.

Oppenheim, R. (1991) 'Cell Death During Development Of The Nervous System', *Annual Review of Neuroscience*, 14, pp. 453–501. doi: 10.1146/annurev.neuro.14.1.453.

Oudet, P., Gross-Bellard, M. and Chambon, P. (1975) 'Electron microscopic and biochemical evidence that chromatin structure is a repeating unit', *Cell*, 4(4), pp. 281–300. doi: 10.1016/0092-8674(75)90149-X.

Paix, A. *et al.* (2015) 'High Efficiency, Homology-Directed Genome Editing in *Caenorhabditis elegans* Using CRISPR/Cas9 Ribonucleoprotein Complexes.', *Genetics*, 201(September), p. genetics.115.179382-. doi: 10.1534/genetics.115.179382.

Park, J.-E. *et al.* (2011) 'Dicer recognizes the 5' end of RNA for efficient and accurate processing', *Nature*, 475(7355), pp. 201–205. doi: 10.1038/nature10198.

Passarge, E. (1979) 'Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago.', *American journal of human genetics*. Elsevier, 31(2), pp. 106–15. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/377956> (Accessed: 4 September 2017).

del Peso, L. *et al.* (2000) 'Disruption of the CED-9/CED-4 Complex by EGL-1 is a Critical Step for Programmed Cell Death in *C. elegans*', *Journal of Biological Chemistry*, 275(35), pp. 27205–11. doi: 10.1074/jbc.M000858200.

del Peso, L., González, V. M. and Núñez, G. (1998) '*Caenorhabditis elegans* EGL-1

Disrupts the Interaction of CED-9 with CED-4 and Promotes CED-3 Activation', *Journal of Biological Chemistry*, 273(50). doi: 10.1074/jbc.273.50.33495.

Petersén A *et al.* (2001) 'Expanded CAG repeats in exon 1 of the Huntington's disease gene stimulate dopamine-mediated striatal neuron autophagy and degeneration.', *Human molecular genetics*, 10(12), pp. 1243–54. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11406606> (Accessed: 14 August 2017).

Plagens, A. *et al.* (2012) 'Characterization of the CRISPR/Cas Subtype I-A System of the Hyperthermophilic Crenarchaeon *Thermoproteus tenax*', *Journal of Bacteriology*, 194(10), pp. 2491–2500. doi: 10.1128/JB.00206-12.

Potts, M. B., Wang, D. P. and Cameron, S. (2009) 'Trithorax, Hox, and TALE-class homeodomain proteins ensure cell survival through repression of the BH3-only gene *egl-1*', *Developmental Biology*, 329(2), pp. 374–385. doi: 10.1016/j.ydbio.2009.02.022.

Qi, S. *et al.* (2010) 'Crystal structure of the *Caenorhabditis elegans* apoptosome reveals an octameric assembly of CED-4.', *Cell*, 141(3), pp. 446–57. doi: 10.1016/j.cell.2010.03.017.

Raffo, A. J. *et al.* (1995) 'Overexpression of *bcl-2* protects prostate cancer cells from apoptosis in vitro and confers resistance to androgen depletion in vivo.', *Cancer research*, 55(19), pp. 4438–45. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7671257> (Accessed: 14 August 2017).

Raj, A. *et al.* (2008) 'Imaging individual mRNA molecules using multiple singly labeled probes', *Nature Methods*, 5(10), pp. 877–879. doi: 10.1038/nmeth.1253.

Raj, A. and Tyagi, S. (2010) 'Detection of Individual Endogenous RNA Transcripts In Situ Using Multiple Singly Labeled Probes', in *Methods in enzymology*, pp. 365–386. doi: 10.1016/S0076-6879(10)72004-8.

Ratner, H. K., Sampson, T. R. and Weiss, D. S. (2016) 'Overview of CRISPR-Cas9 biology', *Cold Spring Harbor Protocols*. Cold Spring Harbor Laboratory Press, 2016(12), pp. 1023–1038. doi: 10.1101/pdb.top088849.

Reddien, P. W., Cameron, S. and Horvitz, H. R. (2001) 'Phagocytosis promotes programmed cell death in *C. elegans*', *Nature*, 412(6843), pp. 198–202. doi: 10.1038/35084096.

Reinke, A. W. *et al.* (2017) 'In vivo mapping of tissue- and subcellular-specific proteomes in *Caenorhabditis elegans*', *Science Advances*. American Association for the

Advancement of Science, 3(5). doi: 10.1126/sciadv.1602426.

Reznik, B. and Lykke-Andersen, J. (2010) 'Regulated and quality-control mRNA turnover pathways in eukaryotes: Figure 1', *Biochemical Society Transactions*, 38(6), pp. 1506–1510. doi: 10.1042/BST0381506.

Richmond, T. J. and Davey, C. A. (2003) 'The structure of DNA in the nucleosome core', *Nature*, 423(6936), pp. 145–150. doi: 10.1038/nature01595.

Rieber, M. S. and Rieber, M. (1994) 'UV radiation induces DNA fragmentation and cell death in B16 melanoma sensitized by bromodeoxyuridine: impaired c-jun induction and defective tyrosine phosphorylation signalling.', *Biochemical and biophysical research communications*, 203(3), pp. 1629–37. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7524489> (Accessed: 18 August 2017).

Ross, J. (1995) 'mRNA stability in mammalian cells.', *Microbiological reviews*, 59(3), pp. 423–50. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7565413> (Accessed: 10 February 2019).

Saetrom, P. *et al.* (2007) 'Distance constraints between microRNA target sites dictate efficacy and cooperativity.', *Nucleic acids research*. Oxford University Press, 35(7), pp. 2333–42. doi: 10.1093/nar/gkm133.

Sandberg, R. *et al.* (2008) 'Proliferating Cells Express mRNAs with Shortened 3' Untranslated Regions and Fewer MicroRNA Target Sites', *Science*, 320(5883), pp. 1643–1647. doi: 10.1126/science.1155390.

Sawa, H., Kouike, H. and Okano, H. (2000) 'Components of the SWI/SNF complex are required for asymmetric cell division in *C. elegans*.', *Molecular cell*, 6(3), pp. 617–24. doi: 10.1016/S1097-2765(00)00060-5.

Schindelin, J. *et al.* (2012) 'Fiji: an open-source platform for biological-image analysis', *Nature Methods*, 9(7), pp. 676–682. doi: 10.1038/nmeth.2019.

Schnabel, H. and Schnabel, R. (1990) 'An Organ-Specific Differentiation Gene, pha-1, from *Caenorhabditis elegans*', *Science*, 250(4981), pp. 686–688. doi: 10.1126/science.250.4981.686.

Schnabel, R. *et al.* (1997) 'Assessing Normal Embryogenesis in *Caenorhabditis elegans* Using a 4D Microscope: Variability of Development and Regional Specification', *Developmental Biology*, 184(2), pp. 234–265. doi: 10.1006/dbio.1997.8509.

Schnabel, R. *et al.* (2006) 'Global cell sorting in the *C. elegans* embryo defines a new mechanism for pattern formation', *Developmental Biology*, 294(2), pp. 418–431. doi: 10.1016/j.ydbio.2006.03.004.

Schwartz, H. T. (2007) 'A protocol describing pharynx counts and a review of other assays of apoptotic cell death in the nematode worm *Caenorhabditis elegans*', *Nature Protocols*, 2(3), pp. 705–714. doi: 10.1038/nprot.2007.93.

Seshagiri, S. and Miller, L. K. (1997) 'Caenorhabditis elegans CED-4 stimulates CED-3 processing and CED-3-induced', *Current Biology*, 7(7), pp. 455–460. doi: 10.1016/S0960-9822(06)00216-8.

Shaham, S. *et al.* (1999) 'Mutational analysis of the *Caenorhabditis elegans* cell-death gene *ced-3*', *Genetics*, 153(4), pp. 1655–1671.

Shatkin, A. J. and Manley, J. L. (2000) 'The ends of the affair: capping and polyadenylation.', *Nature Structural Biology*, 7(10), pp. 838–842. doi: 10.1038/79583.

Shaye, D. D. and Greenwald, I. (2011) 'OrthoList: a compendium of *C. elegans* genes with human orthologs.', *PloS one*. Edited by K. M. Iijima, 6(5), p. e20085. doi: 10.1371/journal.pone.0020085.

Sherrard, R. *et al.* (2017) 'miRNAs cooperate in apoptosis regulation during *C. elegans* development.', *Genes and Development*. Cold Spring Harbor Laboratory Press, 31(2), pp. 209–222. doi: 10.1101/gad.288555.116.

Shimotohno, K. *et al.* (1977) 'Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 74(7), pp. 2734–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/197518> (Accessed: 30 December 2018).

Siepel, A. *et al.* (2005) 'Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes', *Genome Research*, 15(8), pp. 1034–1050. doi: 10.1101/gr.3715005.

Solari, F. and Ahringer, J. (2000) 'NURD-complex genes antagonise Ras-induced vulval development in *Caenorhabditis elegans*', *Current Biology*, 10(4), pp. 223–226. doi: 10.1016/S0960-9822(00)00343-2.

Song, J.-J. *et al.* (2004) 'Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity', *Science*, 305(5689), pp. 1434–1437. doi: 10.1126/science.1102514.

Sorek, R., Kunin, V. and Hugenholtz, P. (2008) 'CRISPR — a widespread system that provides acquired resistance against phages in bacteria and archaea', *Nature Reviews Microbiology*, 6(3), pp. 181–186. doi: 10.1038/nrmicro1793.

Spector, M. S. *et al.* (1997) 'Interaction between the *C. elegans* cell-death regulators CED-9 and CED-4', *Nature*, 385(6617), pp. 653–656. doi: 10.1038/385653a0.

Stamm, S. *et al.* (2005) 'Function of alternative splicing', *Gene*, 344, pp. 1–20. doi: 10.1016/j.gene.2004.10.022.

Stanfield, G. M. and Horvitz, H. R. (2000) 'The *ced-8* gene controls the timing of programmed cell deaths in *C. elegans*.', *Molecular cell*, 5(3), pp. 423–33. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10882128> (Accessed: 30 January 2019).

Stansel, R. M., de Lange, T. and Griffith, J. D. (2001) 'T-loop assembly in vitro involves binding of TRF2 near the 3' telomeric overhang.', *The EMBO journal*. European Molecular Biology Organization, 20(19), pp. 5532–40. doi: 10.1093/emboj/20.19.5532.

Stern, A. *et al.* (2010) 'Self-targeting by CRISPR: gene regulation or autoimmunity?', *Trends in Genetics*, 26(8), pp. 335–340. doi: 10.1016/j.tig.2010.05.008.

Stern, A. and Sorek, R. (2011) 'The phage-host arms race: shaping the evolution of microbes.', *BioEssays: news and reviews in molecular, cellular and developmental biology*. NIH Public Access, 33(1), pp. 43–51. doi: 10.1002/bies.201000071.

Stoeckius, M. *et al.* (2009) 'Large scale sorting of *C. elegans* embryos reveals the dynamics of small RNA expression', *Nature methods*, 6(10), pp. 745–751. doi: 10.1038/nmeth.1370.Large.

Subasic, D. *et al.* (2015) 'Cooperative target mRNA destabilization and translation inhibition by miR-58 microRNA family in *C. elegans*.', *Genome research*. Cold Spring Harbor Laboratory Press, 25(11), pp. 1680–91. doi: 10.1101/gr.183160.114.

Sugimoto, A. *et al.* (2001) 'Many genomic regions are required for normal embryonic programmed cell death in *Caenorhabditis elegans*.', *Genetics*, 158(1), pp. 237–52. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11333233> (Accessed: 17 February 2019).

Sulston, J. E. (1976) 'Post-embryonic development in the ventral cord of *Caenorhabditis elegans*.', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 275(938), pp. 287–97. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8804> (Accessed: 14 August 2017).

Sulston, J. E. *et al.* (1983) 'The embryonic cell lineage of the nematode *Caenorhabditis elegans*.', *Developmental biology*, 100(1), pp. 64–119. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/6684600> (Accessed: 3 January 2019).

Sulston, J. E. and Horvitz, H. R. (1977) 'Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*.', 56(1), pp. 110–56. doi: 10.1016/0012-1606(77)90158-0.

Suzuki, J. *et al.* (2013) 'Xk-Related Protein 8 and CED-8 Promote Phosphatidylserine Exposure in Apoptotic Cells', *Science*, 341(6144), pp. 403–406. doi: 10.1126/science.1236758.

Swarts, D. C. *et al.* (2012) 'CRISPR Interference Directs Strand Specific Spacer Acquisition', *PLoS ONE*. Edited by I. Mokrousov, 7(4), p. e35888. doi: 10.1371/journal.pone.0035888.

Tabara, H. *et al.* (1999) 'The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*.', *Cell*, 99(2), pp. 123–32. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10535731> (Accessed: 30 August 2017).

Terns, M. P. and Terns, R. M. (2011) 'CRISPR-based adaptive immune systems.', *Current opinion in microbiology*, 14(3), pp. 321–7. doi: 10.1016/j.mib.2011.03.005.

Thellmann, M., Hatzold, J. and Conradt, B. (2003) 'The Snail-like CES-1 protein of *C. elegans* can block the expression of the BH3-only cell-death activator gene *egl-1* by antagonizing the function of bHLH proteins', *Development*, 130(17), pp. 4057–4071. doi: 10.1242/dev.00597.

Tock, M. R. and Dryden, D. T. (2005) 'The biology of restriction and anti-restriction', *Current Opinion in Microbiology*, 8(4), pp. 466–472. doi: 10.1016/j.mib.2005.06.003.

Tran, A. T. *et al.* (2019) 'MiR-35 buffers apoptosis thresholds in the *C. elegans* germline by antagonizing both MAPK and core apoptosis pathways', *Cell Death & Differentiation*. Nature Publishing Group, pp. 1–15. doi: 10.1038/s41418-019-0325-6.

Tremethick, D. J. (2007) 'Higher-Order Structures of Chromatin: The Elusive 30 nm Fiber', *Cell*, 128(4), pp. 651–654. doi: 10.1016/j.cell.2007.02.008.

Trent, C., Tsuing, N. and Horvitz, H. R. (1983) 'Egg-laying defective mutants of the nematode *Caenorhabditis elegans*.', *Genetics*. Genetics Society of America, 104(4), pp. 619–47. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11813735> (Accessed: 27 January 2019).

Trojer, P. and Reinberg, D. (2007) 'Facultative Heterochromatin: Is There a Distinctive Molecular Signature?', *Molecular Cell*, 28(1), pp. 1–13. doi: 10.1016/j.molcel.2007.09.011.

Vanfleteren, J. R., van Bun, S. M. and Van Beeumen, J. J. (1987) 'The primary structure of histone H4 from the nematode *Caenorhabditis elegans*', *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 87(4), pp. 847–849. doi: 10.1016/0305-0491(87)90400-7.

Vanfleteren, J. R., Van Bun, S. M. and Van Beeumen, J. J. (1987) 'The primary structure of histone H3 from the nematode *Caenorhabditis elegans*', *FEBS Letters*, 211(1), pp. 59–63. doi: 10.1016/0014-5793(87)81274-7.

Vanfleteren, J. R., Van Bunt, S. M. and Van Beeumen, J. J. (1989) 'The histones of *Caenorhabditis elegans*: no evidence of stage-specific isoforms', *FEBS Letters*, 257(2), pp. 233–237. doi: 10.1016/0014-5793(89)81541-8.

Ventura, A. *et al.* (2008) 'Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters.', *Cell*. Howard Hughes Medical Institute, 132(5), pp. 875–86. doi: 10.1016/j.cell.2008.02.019.

Visa, N. *et al.* (1996) 'A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export.', *The Journal of cell biology*. Rockefeller University Press, 133(1), pp. 5–14. doi: 10.1083/JCB.133.1.5.

Vogt, C. (1842) *Untersuchungen über die Entwicklungsgeschichte der Geburtshelferkröte (Alytes obstetricans)*. Jent und Gassman, Solothurn.

Wahle, E. and Keller, W. (1992) 'The Biochemistry of 3'-End Cleavage and Polyadenylation of Messenger RNA Precursors', *Annual Review of Biochemistry*. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA , 61(1), pp. 419–438. doi: 10.1146/annurev.bi.61.070192.002223.

Wang, J. *et al.* (2015) 'The *C. elegans* COE transcription factor UNC-3 activates lineage-specific apoptosis and affects neurite growth in the RID lineage', *Development*, 142(March), pp. 1–11. doi: 10.1242/dev.119479.

Wang, X. *et al.* (2002) 'Mechanisms of AIF-Mediated Apoptotic DNA Degradation in *Caenorhabditis elegans*', *Science*, 298(5598), pp. 1587–1592. doi: 10.1126/science.1076194.

Ward, J. D. (2015) 'Rapid and Precise Engineering of the *Caenorhabditis elegans* Genome with Lethal Mutation', *Genetics*. Genetics Society of America, 199(February), pp. 363–377. doi: 10.1534/genetics.114.172361.

Wei, Y., Terns, R. M. and Terns, M. P. (2015) 'Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation.', *Genes & development*. Cold Spring Harbor Laboratory Press, 29(4), pp. 356–61. doi: 10.1101/gad.257550.114.

Wickens, M. (1990) 'In the beginning is the end: regulation of poly(A) addition and removal during early development.', *Trends in biochemical sciences*, 15(8), pp. 320–4. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2204159> (Accessed: 10 February 2019).

Wicky, C. *et al.* (1996) 'Telomeric repeats (TTAGGC)_n are sufficient for chromosome capping function in *Caenorhabditis elegans*.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 93(August), pp. 8983–8988. doi: 10.1073/pnas.93.17.8983.

Wightman, B. *et al.* (1991) 'Negative regulatory sequences in the *lin-14* 3'-untranslated region are necessary to generate a temporal switch during *Caenorhabditis elegans* development.', *Genes & development*, 5(10), pp. 1813–24. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1916264> (Accessed: 19 August 2017).

Wightman, B., Ha, I. and Ruvkun, G. (1993) 'Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*.', *Cell*, 75(5), pp. 855–62. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8252622> (Accessed: 19 August 2017).

Wigler, M. H. and Axel, R. (1976) 'Nucleosomes in metaphase chromosomes.', *Nucleic acids research*, 3(6), pp. 1463–71. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/958895> (Accessed: 2 September 2017).

Will, C. L. and Lührmann, R. (2011) 'Spliceosome structure and function.', *Cold Spring Harbor perspectives in biology*. Cold Spring Harbor Laboratory Press, 3(7). doi: 10.1101/cshperspect.a003707.

Winn, J. *et al.* (2011) 'Hox and a Newly Identified E2F Co-repress Cell Death in *Caenorhabditis elegans*', *Genetics*, 188(4). Available at: http://www.genetics.org/content/188/4/897?ijkey=62cb92e43da53a086fc6911974eb932f39fc4690&keytype=tf_ipsecsha (Accessed: 30 August 2017).

Winter, J. *et al.* (2009) 'Many roads to maturity: MicroRNA biogenesis pathways and their regulation', *Nature Cell Biology*. Nature Publishing Group, pp. 228–234. doi: 10.1038/ncb0309-228.

Woo, J. S. *et al.* (2003) 'Unique structural features of a BCL-2 family protein CED-9 and biophysical characterization of CED-9/EGL-1 interactions', *Cell Death and Differentiation*. Cell Death Differ, 10(12), pp. 1310–1319. doi: 10.1038/sj.cdd.4401303.

Wood, W. B. (1988) *The nematode C. elegans*. Edited by W. B. Wood. Cold Spring Harbor Laboratory Press, New York.

Woodcock, C. L. and Dimitrov, S. (2001) 'Higher-order structure of chromatin and chromosomes.', *Current opinion in genetics & development*, 11(2), pp. 130–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11250134> (Accessed: 2 September 2017).

Woodcock, C. L. and Ghosh, R. P. (2010) 'Chromatin higher-order structure and dynamics.', *Cold Spring Harbor perspectives in biology*. Cold Spring Harbor Laboratory Press, 2(5), p. a000596. doi: 10.1101/cshperspect.a000596.

Wormatlas.org (2017) *CAENORHABDITIS elegans AS A GENETIC ORGANISM*. Available at: <http://www.wormatlas.org/ver1/handbook/anatomyintro/anatomyintro.htm#Lifecycle> (Accessed: 14 August 2017).

Wormbase.org (2017a) *mir-35 (gene) - WormBase: Nematode Information Resource*. Available at: http://www.wormbase.org/species/c_elegans/gene/WBGene00003263#0-9gc1-10 (Accessed: 22 October 2017).

Wormbase.org (2017b) *mir-42 (gene) - WormBase: Nematode Information Resource*. Available at: http://www.wormbase.org/species/c_elegans/gene/WBGene00003270#0-9g-3 (Accessed: 22 October 2017).

Wu, D. *et al.* (1997) 'Interaction and Regulation of the Caenorhabditis elegans Death Protease CED-3 by CED-4 and CED-9', *Journal of Biological Chemistry*, 272(34). doi: 10.1074/jbc.272.34.21449.

Wu, D. *et al.* (1999) 'C. elegans MAC-1, an essential member of the AAA family of ATPases, can bind CED-4 and prevent cell death.', *Development (Cambridge, England)*, 126(9), pp. 2021–31. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10101135>

(Accessed: 31 July 2019).

Wu, E. *et al.* (2010) 'Pervasive and cooperative deadenylation of 3'UTRs by embryonic microRNA families.', *Molecular cell*. Elsevier Inc., 40(4), pp. 558–70. doi: 10.1016/j.molcel.2010.11.003.

Wu, Q. *et al.* (2017) 'Coal combustion related fine particulate matter (PM2.5) induces toxicity in *Caenorhabditis elegans* by dysregulating microRNA expression.', *Toxicology research*. Royal Society of Chemistry, 6(4), pp. 432–441. doi: 10.1039/c7tx00107j.

Wu, Y. *et al.* (2013) 'Correction of a Genetic Disease in Mouse via Use of CRISPR-Cas9', *Cell Stem Cell*, 13(6), pp. 659–662. doi: 10.1016/j.stem.2013.10.016.

Xia, J., Bogardus, C. and Prochazka, M. (1999) 'A Type 2 Diabetes-Associated Polymorphic ARE Motif Affecting Expression of PPP1R3 Is Involved in RNA–Protein Interactions', *Molecular Genetics and Metabolism*, 68(1), pp. 48–55. doi: 10.1006/mgme.1999.2884.

Xia, Z. *et al.* (2014) 'Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types', *Nature Communications*, 5(1), p. 5274. doi: 10.1038/ncomms6274.

Xie, X. *et al.* (2005) 'Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals', *Nature*, 434(7031), pp. 338–345. doi: 10.1038/nature03441.

Yan, N. *et al.* (2004) 'Structural, Biochemical, and Functional Analyses of CED-9 Recognition by the Proapoptotic Proteins EGL-1 and CED-4', *Molecular Cell*, 15(6), pp. 999–1006. doi: 10.1016/j.molcel.2004.08.022.

Yang, X. (1998) 'Essential Role of CED-4 Oligomerization in CED-3 Activation and Apoptosis', *Science*, 281(5381), pp. 1355–1357. doi: 10.1126/science.281.5381.1355.

Yi, R. *et al.* (2003) 'Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs', *Genes & Development*, 17(24), pp. 3011–3016. doi: 10.1101/gad.1158803.

Yuan, J. and Horvitz, H. R. (1992) 'The *Caenorhabditis elegans* cell death gene *ced-4* encodes a novel protein and is expressed during the period of extensive programmed cell death.', *Development (Cambridge, England)*, 116(2), pp. 309–20. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1286611> (Accessed: 30 January 2019).

Zarkower, D. (2006) 'Somatic sex determination', *WormBook*. doi: 10.1895/wormbook.1.84.1.

Zarkower, D. and Hodgkin, J. (1993) 'Zinc fingers in sex determination: only one of the two *C. elegans* Tra-1 proteins binds DNA in vitro.', *Nucleic acids research*, 21(16), pp. 3691–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8367286> (Accessed: 27 January 2019).

Zeng, Y. and Cullen, B. R. (2004) 'Structural requirements for pre-microRNA binding and nuclear export by Exportin 5', *Nucleic Acids Research*, 32(16), pp. 4776–4785. doi: 10.1093/nar/gkh824.

Zhang, J., Tsapralis, G. and Bowden, G. T. (2008) 'Nucleolin Stabilizes Bcl-X_L Messenger RNA in Response to UVA Irradiation', *Cancer Research*, 68(4), pp. 1046–1054. doi: 10.1158/0008-5472.CAN-07-1927.

6 List of figures and tables

Figure 1. Representative images of <i>C. elegans</i>	5
Figure 2. Phases of apoptosis in <i>C. elegans</i>	9
Figure 3. Representation of the central cell death pathway in <i>C. elegans</i>	11
Figure 4. Regulation of <i>egl-1</i> transcription.....	13
Figure 5. Biogenesis of microRNAs.....	23
Figure 6. Labeling targeted DNA sequence with a dCas9::GFP::GFP/sgRNA complex.	29
Figure 7. <i>egl-1</i> 3'UTR harbors conserved micro RNA binding sites	35
Figure 8. Distinct waves of cell death occur during AB lineage development in <i>C. elegans</i>	37
Figure 9. Embryos lacking miR-35 family miRNAs exhibit a large cell corpse phenotype.....	40
Figure 10. Embryonic lethality is not recovered by blocking apoptosis in miR-35 family mutants.....	42
Figure 11. The 3'UTR of <i>egl-1</i> is a target of miR-35 and miR-58 family miRNAs <i>in vivo</i>	45
Figure 12. The <i>egl-1</i> 3'UTR causes a reduction in mRNA copy number of the <i>P_{mai-2} gfp::h2b</i> reporter.	46
Figure 13: the <i>egl-1</i> 3'UTR with inserted repair template after CRISPR/Cas9 induced NHEJ.	51
Figure 14: The donor ssDNA integration changes the predicted <i>egl-1</i> 3'UTR folding.	52
Figure 15: Extruded cells in embryos carrying the <i>egl-1</i> (<i>bc275</i>) variant.....	53
Figure 16: Representative smRNA FISH images of <i>egl-1</i> mRNA in the RID lineage	56
Figure 17: Quantification of <i>egl-1</i> mRNA copy number in the RID lineage.....	57
Figure 18. Labeling targeted DNA sequence with a dCas9::GFP::GFP.	59

Figure 19. Targeting telomeric repeats with dCas9::GFP::GFP/sgRNA complex..	62
Figure 20: <i>egl-1</i> expression level (TPM) during embryonic development.....	69
Figure 21: Ten representative candidate genes with best fit (eight points) to the defined expression pattern.....	71
Figure 22: GO enrichment analysis visualization	75
Figure 23. Genetic model of miR-35 and miR-58 family dependent control on <i>egl-1</i> BH3-only in the RID lineage.....	82

Table 1: Inappropriate cell death in <i>mir-35</i> family mutants is enhanced upon loss of the <i>mir-58</i> family	39
Table 2. miR-35 family miRNAs act upstream or in parallel to <i>egl-1</i>	43
Table 3. Abnormal cell death in <i>miR-35</i> family mutants affects mothers and sisters of programmed cell death.	48
Table 4: co-CRISPR approach utilizing <i>dpy-10</i> and <i>pha-1</i> selectable marker for <i>egl-1</i> 3'UTR modification.....	50
Table 5: Extra cells in the anterior pharynx.....	54
Table 6: <i>C. elegans</i> injected for <i>mos1</i> -mediated single copy integration of the <i>dCas9::gfp::gfp</i> fusion construct.....	60
Table 7. Collection of candidate genes via Gene Ontology by keyword.....	65
Table 8: Genes not identified based on keyword-based search within the <i>C. elegans</i> gene ontology.	67
Table 9: Candidate genes are mapped against expression pattern and sorted in categories.	70
Table 10: Overrepresentation analysis for GO term programmed cell death (GO:0012501) overall candidate groups.	73
Table 14: Apoptotic and developmental process related enriched GO terms within the candidate list.....	74
Table 11: Top five enriched GO terms of candidate genes with ideal expression pattern.	76
Table 12: Section of the GO term enrichment analysis of candidate genes with ideal expression pattern.	77
Table 13: Cell death associated candidate genes identified via GO enrichment analysis.....	78
Table 15. Alleles used in this work	96
Table 16. Transgenes, extrachromosomal arrays used in this work.	97

Table 17: Transgenes and Alleles generated	100
Table 18: Injection mix for MosSCI.....	100
Table 19: Sequencing primers for 3'UTR reporter transgenes.....	102
Table 20. Sequences of <i>egl-1</i> smRNA FISH probes.....	103
Table 21. Sequences of <i>unc-3::gfp</i> smRNA FISH probes.....	103
Table 22: <i>pha-1 egl-1</i> co-Conversion injection mix	109
Table 23: <i>dpy-10 egl-1</i> co-Conversion injection mix	109

7 Curriculum vitae

Academic Career

03/2013 – 07/2017	Practical part of PhD thesis in biology Ludwig-Maximilians-Universität Munich, Focus: Cell and Developmental Biology Title: „ Regulation of the central apoptotic pathway during <i>C. elegans</i> development “ Prof. Dr. Barbara Conradt
04/2010 – 11/2012	Master of Biology Ludwig-Maximilians-Universität Munich, Focus: Human biology and Genetics Master thesis: „Establishment of ChIP-trapping assay with Dnmt1 binder“ Prof. Dr. Heinrich Leonhardt
09/2006 – 04/2010	Bachelor of Biology Ludwig-Maximilians-Universität Munich, Focus: Genetics, Mikrobiology
08/2003 – 09/2006	Vocational training as state-certified biologic-technical assistant Lessing Berufskolleg, Düsseldorf
08/2003 – 06/2006	General Abitur Lessing Berufskolleg, Düsseldorf Focus: Biology, Chemistry

Publications

Genome-wide RNAi screen for regulators of UPRmt in *Caenorhabditis elegans* mutants with defects in mitochondrial fusion.
Haeussler S, Yeroslaviz A, Rolland SG, **Luehr S**, Lambie EJ, Conradt B.
G3 (Bethesda). 2021 Mar 30;jkab095. doi: 10.1093/g3journal/jkab095. Epub ahead of print. PMID: 33784383.

miRNAs cooperate in apoptosis regulation during *C. elegans* development
Sherrard R and **Luehr S**, Holzkamp H, McJunkin K, Memar N, Conradt B.
Genes Dev. 2017 Jan 15;31(2):209-222

P-value-based regulatory motif discovery using positional weight matrices.
Hartmann H, Guthöhrlein EW, Siebert M, **Luehr S**, Söding J.
Genome Res. 2013 Jan;23(1):181-94.

The XXmotif web server for eXhaustive, weight matriX-based motif discovery in nucleotide sequences.
Luehr S, Hartmann H, Söding J.
Nucleic Acids Res. 2012 Jul;40(Web Server issue):W104-9