A small RNA pathway functions during the oocyte-to-embryo transition in *Caenorhabditis elegans*

Inauguraldissertation

zur Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Christina Fassnacht

aus Deutschland

Basel, 2017

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel

edoc.unibas.ch

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät der Universität Basel auf Antrag von

Prof. Dr. Susan Gasser

Prof. Dr. Christian Lehner

Dr. Rafal Ciosk

Basel, den 13. Dezember 2016

Prof. Dr. Jörg Schibler

(Dekan der Philosophisch-Naturwissenschaftlichen Fakultät der Universität Basel)

Table of contents

1 Summary

A fundamental reprogramming occurs during the oocyte-to-embryo transition (OET) of an animal, from a transcriptionally silent, differentiated oocyte to a pluripotent early embryo. The two major events contributing to the cell fate change are maternal transcript degradation and embryonic genome activation (EGA). We used *Caenorhabditis elegans* as a model to study the mechanisms controlling EGA during this developmental window and performed a genetic screen for the unbiased identification of new involved players. We discovered several factors inhibiting a precocious onset of EGA in germ cells, among them several components of the CSR-1 endogenous-small interfering RNA (endo-siRNA) pathway. This pathway employs the Argonaute protein CSR-1 to load small RNAs and has been implicated in diverse processes in germ cells and early embryos. Here, we demonstrate a role for this pathway in the repression of embryonic genes in developing oocytes. Moreover, we found that the catalytic activity of CSR-1 to degrade RNAs is required for this function. The importance of such a post-transcriptional role is in accordance with previous findings, which described a function for the RNA-binding proteins GLD-1 and LIN-41 in germ cell fate control. Our discovery of a direct interaction of CSR-1 and LIN-41 suggests a co-regulatory role for the two proteins in EGA repression. Together, the results contribute to the understanding of the reprogramming during the *C. elegans* OET and highlight for the first time the importance of endo-siRNAs in this process.

2 Introduction

2.1 Cell fate maintenance and reprogramming during the oocyte-to-embryo transition

This chapter is adapted from the book chapter of the same name, published in:

Signaling-Mediated Control of Cell Division

From Oogenesis to Oocyte-to-Embryo Development

© Springer, 2017

Authors: Christina Fassnacht and Rafal Ciosk

2.1.1 The oocyte-to-embryo transition

Cellular differentiation leads to the formation of highly specialized cells, such as neurons and muscles, which fulfill complex functions. Cellular reprogramming is the opposite process, where a specialized cell, with a restricted developmental potential, is turned into a pluripotent cell that can give rise to different types of cells. During development, a complete reprogramming into pluripotency is restricted to the oocyte-to-embryo transition (OET). During this transition, a fertilized oocyte develops into a pluripotent embryo, which has the potential to give rise to an entire new individual. The main cellular processes occurring during the OET are oocyte maturation, ovulation, and fertilization. Those processes will be described in more detail for the model *Caenorhabditis elegans* in chapter 2.2. In this chapter, we focus on the changes in the transcriptome, underlying the transition from maternal to embryonic control of development. In addition, we concentrate on mechanisms operating in the oocytes, because unfertilized oocytes display a broad developmental potential, which is evident by their ability to give rise to developmentally plastic parthenogenetic cell lines (Brevini et al. 2012), or even offspring when paternal imprinting is mimicked (Kono et al. 2004). The initial stages of the OET, from oocyte maturation to early embryogenesis, occur in the absence of *de novo* Pol II-mediated transcription and are controlled, post-transcriptionally, until the embryo begins producing its own transcripts during EGA (LaMarca et al. 1975; Moore et al. 1974). Our current understanding of the mechanisms controlling developmental reprogramming during the OET remains very limited. Experiments with pluripotent embryonic stem (ES)

cells have shown that pluripotency can be regulated at each step of gene expression. However, which of the mechanisms described in ES cells are employed during development remains to be, to a large extent, determined. During development, a successful OET depends on the ability to maintain the germ cell fate in a developing oocyte, which includes the maintenance of germ line-specific gene expression and the inhibition of soma/embryo-specific expression. Failures in this process can result in a precocious acquisition of pluripotency and somatic-like, teratomatous differentiation. In an early embryo, the regulation of gene expression is reversed, resulting in the activation of somatic/embryonic genes and the inhibition of germline genes. In this chapter, we will discuss critical factors and emerging mechanisms controlling this dramatic reprograming of gene expression.

2.1.2 Developmental potential of germ cells

Germ cells are typically the only cells giving rise to the offspring. This striking property of germ cells led to the formulation of the germ-plasm theory of heredity, also known as the Weismann theory of heredity, which states that only germ cells transmit genetic information to the next generation (Weismann 1893). Indeed, the development and maintenance of germ cells is crucial for the survival of species with sexual reproduction and germ cell-specific genome surveillance mechanisms have evolved to ensure genome integrity. Germ cells originate from primordial germ cells (PGCs) that are usually specified during embryonic development, for a recent review on the specification of germ cell fate in different organisms see (Strome and Updike 2015). PGCs and also adult germ cells have the potential to give rise to other types of cells, as evident by the different types of pluripotent cell lines derived from germ cells at different stages of development (Kerr et al. 2006). For example, embryonic germ (EG) cells can be derived from mouse PGCs between the embryonic days E8.5 until E13.5 (Cantone and Fisher 2013). The EG cells are in many aspects like blastocyst-derived embryonic stem (ES) cells, in that they can generate chimeras and display germline transmission (Smith 2001). The pluripotency of germ cells is also manifested by the properties of a specific germ cell tumor (GCT) called teratoma. In human teratomas, germ cells abnormally differentiate into somatic cell types, such as neurons, muscles, hair or teeth (Ulbright 2005). They can arise in ovary and testis, with testicular GCTs accounting for the most common solid malignancy found in males between the ages of 15 and 40 years (Hussain et al. 2008). Origins of GCT development during human and mouse gametogenesis have been recently reviewed in (Dolci et al. 2015). They include epigenetic remodeling, as well as defects in proliferation, apoptosis and the mitotic to meiotic switch. Undifferentiated cells from teratomas can also be grown *ex vivo* as embryonic carcinoma (EC) cells and can be differentiated into cells of all three germ layers in culture (Smith 2001). The formation of teratomas

shows that, to ensure normal development, there need to be repressive mechanisms operating in adult germ cells to control pluripotency. It is mostly the work from the nematode *C. elegans* that contributed to the identification of pluripotency repressors, which will be discussed later. A final evidence for the underlying pluripotency of germ cells is the ability of the oocyte cytoplasm to reprogram somatic nuclei. This was experimentally demonstrated, by using the cytoplasm of an enucleated *Xenopus* oocyte to reprogram an intestine nucleus (Gurdon and Uehlinger 1966). Since then, reprogramming has been mostly studied using cell culture. This revealed some important regulators, most notably the pluripotency transcription factors Oct4, Sox2, Klf4 and c-Myc (Takahashi and Yamanaka 2006). However, the reprogramming remains a very inefficient process, indicating that repressive mechanisms stabilizing the differentiated state may exist (Pasque et al. 2011). Arguably, understanding the mechanisms controlling the natural reprogramming during OET may be useful for improving the *in vitro* reprogramming techniques.

2.1.3 Control of embryonic genome activation

Two major processes occur during the OET to reprogram the transcriptome. One is the maternal clearance, the degradation of many maternal mRNAs and proteins deposited into the oocyte. The mechanisms behind the maternal clearance have been recently reviewed (Barckmann and Simonelig 2013; Walser and Lipshitz 2011). The other process is the embryonic genome activation (EGA), the onset of transcription in the early embryo. In the following, we concentrate on the advances that have been made in understanding the timing, dynamics and regulation of EGA, using different model organisms. Embryonic transcription initiates at a species-specific time, which ranges in vertebrates from 1 to 2 cell cycles in mouse, to cell cycles 6-9 in *Xenopus* and zebrafish embryos (Lee et al. 2014). Though, from an absolute time perspective EGA occurs rather later in the mouse, as the first cell division occurs a long time after fertilization. Among invertebrates, embryonic transcription starts at cell cycle 2 in the nematode *C. elegans* and at cell cycle 6 in the fruit fly *D. melanogaster* (Lee et al. 2014). A common sequence in EGA across model organisms is the activation of transcription in two waves. In the first wave only a few embryonic mRNAs are transcribed, whereas in the second, later wave, the bulk of transcription occurs. The dynamics of these two waves vary between organisms. In mouse and in *C. elegans*, the second wave of transcription rapidly proceeds the first, whereas in the fruit fly there is a long pause between the waves, with the first at cell cycle 6 and the major wave only at cell cycle 13, where the embryo consists already of roughly 6000 cells (Lee et al. 2014).

Different models have been proposed to explain the regulation of EGA, including the nucleocytoplasmic (N/C) ratio model, the maternal clock model and the transcript abortion model. The N/C ratio model postulates that for EGA to begin, an EGA repressor in the cytoplasm must be titrated away, as the nuclear to cytoplasmic ratio increases during cell divisions (Newport and Kirschner 1982). One candidate repressor to be regulated by titration is xDnmt1 in *Xenopus* early embryos, whose molecular function will be discussed later. The maternal clock model, in contrast, is cell-cycle independent and hypothesizes that the accumulation or increase in activity of an EGA activator to a certain threshold is needed to start EGA, rather than the titration of a repressor (Lee et al. 2014). One example for the regulation by a maternal clock is TAF-4, a *C. elegans* TFIID subunit, whose activity is controlled by a fertilization-dependent mechanism. The transcript abortion model proposes that the properties of the cell cycle control early gene expression. In many organisms the first embryonic divisions occur very fast, cycling only between S and M phases and lacking the gap phases of the cell cycle. The transcript abortion model postulates that during such fast cell cycles, zygotic transcripts are aborted and can only be fully transcribed when the cell cycles lengthen (Tadros and Lipshitz 2009). The abortion of a nascent long transcript was, indeed, observed during early divisions in *Drosophila* (Shermoen and O'Farrell 1991). Further experimental evidence comes from studies, where a block of the cell cycle leads to premature EGA in *Xenopus* (Kimelman et al. 1987) and *Drosophila* embryos (Edgar and Schubiger 1986). It is tempting to speculate that EGA regulation by the cell cycle is more important in organisms with fast embryonic divisions, like *Drosophila* with early mitotic cycles of 8 minutes only, whereas could be less relevant in organisms with slow embryonic divisions, like mice, where the first two cell cycles take 2 days. However, the short length of early transcribed genes is conserved among species (Heyn et al. 2014), including mice, leaving the questions open if gene length is also restricting gene expression in early mouse embryos. In summary, the three general EGA models are not mutually exclusive and it is possible that a combination of them allows the tight regulation of EGA. The recent identification of critical EGA repressors and activators has shed some light into the molecular mechanisms controlling EGA. When considering the major changes that have to occur from a transcriptionally silent oocyte, with tightly packed DNA, to an early embryo undergoing rapid DNA replication and transcription, it is perhaps not surprising that mechanisms, which globally regulate DNA packaging and transcription, have been identified to be crucial for the reprogramming during OET. In addition, although little is known about post-transcriptional regulation of reprogramming, recent examples highlight the contribution of RNA regulation to a successful OET. Examples for the transcriptional and post-transcriptional regulation of reprogramming will be presented in the following.

2.1.4 Transcriptional regulation of reprogramming

Histone variants

Changes in the chromatin regulate the accessibility of transcription factors to DNA and, thus, transcription. Heterochromatin contains tightly packed DNA with a limited access by the transcription machinery. In contrast, euchromatin is loosely packed, allowing a relatively easy access to DNA. Chromatin is established by the packing of DNA into nucleosomes, and is influenced by nucleosome composition and epigenetic modifications on histone tails of the nucleosomes. Nucleosomes consist of four histone core proteins (H2A, H2B, H3, H4), which can be exchanged by variant histones with unique regulatory properties. Histone exchange, from repressive gamete-specific histones to somatic histones, is likely contributing to the establishment of a permissive chromatin state in embryos. It was shown in mouse that the repressive macroH2A variant in the female pronucleus is progressively lost after fertilization and during EGA onset (Chang et al. 2005). Similarly, studies from *Drosophila* showed that an early embryonic and germline-specific H1 variant, named dBigH1, is exchanged for a somatic version during embryonic development, coinciding with the transcriptional activation (Perez-Montero et al. 2013). Consequently, the loss of dBigH1 leads to a precocious activation of Pol II-dependent transcription in the early embryo (Perez-Montero et al. 2013).

Chromatin remodelers

Another type of chromatin regulation that influences gene expression is the nucleosome positioning, which is mediated by chromatin remodeling complexes. The finding that the remodelers contribute to EGA came from the mouse, as the loss of maternal Brg1 protein, a catalytic subunit of the SWI/SNF chromatin remodeler complex, leads to a reduced transcription of 30 % of embryonic genes and embryonic-arrest (Bultman et al. 2006). However, the question remains whether Brg-1 contributes to a general opening of chromatin or whether Brg1 might be specifically targeted to the promoters of embryonic genes, making them accessible for the transcription machinery.

Epigenetic modifications: Histone marks

As indicated earlier, chromatin can be also influenced by specific modifications of histone tails in the nucleosomes, either promoting or repressing transcription. An enrichment of the activating histone H3K4me3 mark was found on embryonic gene promoters, preferentially on those with house-keeping

roles that are expressed in *Xenopus* and Zebrafish embryos after EGA onset (Akkers et al. 2009; Vastenhouw et al. 2010; Lindeman et al. 2011). In contrast, the repressive histone H3K27me3 mark was present on the promoters of later expressed genes with developmental roles. This suggests a role for these histone modifications in distinguishing earlier vs. later expressed embryonic genes. Another interesting finding is the appearance of these histone marks only at the time of active transcription in the early embryos (Li et al. 2014; Vastenhouw et al. 2010; Akkers et al. 2009), indicating the presence of a relatively unmodified, naïve chromatin state of the embryo before EGA in several species. Therefore, histone marks likely do not contribute to the transcriptional regulation of the first embryonic genes, but rather direct gene expression later during embryonic development.

Epigenetic modifications: DNA methylation

In mammals, also DNA methylation is thought to contribute to EGA regulation. DNA methylation is the addition of a methyl group to cytosines in CpG dinucleotides, and methylation of promoter sequences is thought to have a repressive effect on transcription. Consistent with a repressive function, studies in mouse have reported a global DNA demethylation shortly after fertilization and during EGA, which might contribute to an open pluripotent chromatin state (Paranjpe and Veenstra 2015). However, in other vertebrate species the opposite trend was reported: zebrafish and *Xenopus* embryonic genomes appear to be globally hypermethylated during EGA (Paranjpe and Veenstra 2015). In addition, no obvious correlation between DNA methylation and transcriptional repression of genes seems to exist in *Xenopus* embryos (Bogdanovic et al. 2011). The DNA methyltransferase xDnmt1 was identified in *Xenopus* as an EGA repressor (Stancheva and Meehan 2000), though its silencing function appears to be independent of its catalytic activity and it is suggested that Dnmt1 acts as a direct transcriptional repressor before the onset of EGA (Dunican et al. 2008). In sum, the contribution of DNA methylation dynamics to EGA regulation may vary from species-to-species.

Alternative promoters

An important contribution towards understanding the link between chromatin modifications and transcription comes from studies using cap analysis of gene expression (CAGE) in early zebrafish development to map transcription start sites (TSSs) (Nepal et al. 2013; Haberle et al. 2014). CAGE data showed differences in the TSS usage between maternal and embryonic transcripts and identified an A/T-rich (W-box) motif upstream the maternal TTSs and G/C-rich regions downstream of zygotic TSSs, with the two motifs often co-existing in promoters of constitutively expressed genes. This suggests that specific

sequences, presumably via recruitment of specific transcription factors, direct the dynamic use of maternal and embryonic TSSs during the OET. Interestingly, H3K4me3-marked nucleosomes are positioned specifically at G/C-rich embryonic sequences, which provides a link between chromatin remodeling and sequence information. It remains to be determined which transcription factors bind these motifs. The use of alternative promoters during the OET has also been reported from a study in mouse, where retrotransposons that are highly expressed in mature oocytes and early embryos, can serve as alternative promoters and first exons for embryonic genes (Peaston et al. 2004). This results in the production of early embryo-specific chimeric transcripts. The production of such chimeric transcripts has been reported for a number of host genes and it remains to be shown how widespread this type of regulation is.

Transcription factors

The availability of both general (pioneering) and more specific transcription factors has been found to contribute to the regulation of embryonic gene expression. In addition, work from different model organisms has shown that the regulation of Pol II initiation and elongation, as well as of components of the pre-initiation complex (PIC), contribute to EGA control (Zurita et al. 2008). The activity of Pol II can be regulated via phosphorylation of serine residues of tandem repeats on the C-terminal domain (CTD). Particularly the phosphorylation of Serines 2 and 5 of the CTD is crucial during the initiation and elongation of gene transcription. Pol II phosphorylation in mouse correlates well with transcription states during the OET. Transcriptionally silent oocytes show decreasing levels of CTD phosphorylation, whereas after fertilization phosphorylation increases again before EGA (Bellier et al. 1997). Another level of Pol II and PIC component regulation is their cellular localization. Pol II and PIC components have been shown to localize to the oocyte cytoplasm in many organisms and only after their nuclear translocation in early embryos, the general transcription machinery becomes active. One example is the sequestration of TAF-4, the *C. elegans* TFIID subunit, by two repressors called OMA-1 and OMA-2 in the cytoplasm of the one- and two-cell embryo (Guven-Ozkan et al. 2008). Responsible for the inhibitory function of OMA-1/-2 is the fertilization-dependent phosphorylation by MBK-2. This phosphorylation ultimately leads to OMA-1/-2 degradation at the four-cell stage and the release of TAF-4 into the nucleus, where assembly of TFIID can occur and transcription begin.

The establishment of an open chromatin state and the activation of the basal transcription machinery during OET provide permissive mechanisms for EGA to occur. However, it is expected that other repressive

and activating transcription factors provide instructive mechanisms to gene expression and mediate fine-tuning of transcription during embryo development. The identification of transcription factors with an intrinsic chromatin remodeling function, like Zelda in *D. melanogaster*, suggests that both regulatory mechanisms can even occur at the same time. Zelda is a maternally deposited transcription factor that binds to a heptamer motif, called TAGteam motif, in early embryonic gene promoters and activates their transcription (Liang et al. 2008; De Renzis et al. 2007; ten Bosch et al. 2006). Loss of Zelda leads to a failure in the activation of over a hundred embryonic genes, and defects in cellularization and pattern formation in the early embryo (Liang et al. 2008). In addition to the activation of these early genes, Zelda associates with the promoters of a thousand genes, whose transcription occurs later in embryonic development, suggesting a broader role for Zelda in licensing genes for transcription (Harrison et al. 2011; Nien et al. 2011). Indeed, Zelda was recently found to mediate chromatin accessibility and can therefore be seen as a pioneer transcription factor, facilitating transcription factor binding genome-wide (Schulz et al. 2015). Also in zebrafish, transcription factors that likely possess pioneering activity have been discovered to contribute to EGA. They include the pluripotency transcription factors Nanog, Pou5f3 (formerly Pou5f1, an Oct4 homolog) and Sox19b (a Sox2 homolog), which have been found before to mediate reprogramming to iPS cells (Takahashi and Yamanaka 2006). Their combined loss leads to a failure in the activation of more than 75% of early embryonic genes and developmental arrest of zebrafish embryos before gastrulation (Lee et al. 2013). The observed correlation between the chromatin association of Pou5f3 and Sox2, and the enrichment in H3K4me3 and Pol II binding, suggest that these transcription factors, like in ES and iPS cells, also function during early development as pioneering factors to facilitate *de novo* gene expression (Leichsenring et al. 2013). In addition, it must be noted that many EGA activators, like Zelda and the pluripotency transcription factors Nanog, Pou5f3 and Sox19b, also contribute to the maternal mRNA clearance, by activating the expression of factors mediating mRNA degradation; for example the expression of microRNA miR-430 is activated by Nanog (Lee et al. 2013). Thus, these transcription factors provide a connection between the mechanisms driving EGA and maternal product degradation.

Concluding, transcriptional regulation of reprogramming plays a critical role during the OET (the involved factors are illustrated in Figure 2.1 and summarized in Table 2.1). The regulation of chromatin and basal transcription machinery provide mechanisms to globally control transcriptional competence in germ cells and early embryos. However open remaining questions are: How is specificity of gene regulation achieved? What marks embryonic genes as active and maternal genes as inactive? What induces the

switch from maternal to embryonic gene expression? What determines the level and spatial regulation of embryonic gene expression? The dynamic use of maternal and embryonic transcription start sites provides one exciting possibility how differential transcription can be achieved. However, the molecular mechanisms underlying the differential expression, for example transcription factors recognizing the motifs, remain to be determined. Also the identification of the TAG team motif in early *Drosophila* genes and the identification of Zelda, as the associating TF, provided some insights into embryonic gene regulation. However, such motifs in embryonic gene promoters have not been identified in other model organisms. This raises the question how widespread and conserved the utilization of embryo-specific TF-binding motifs is across organisms.

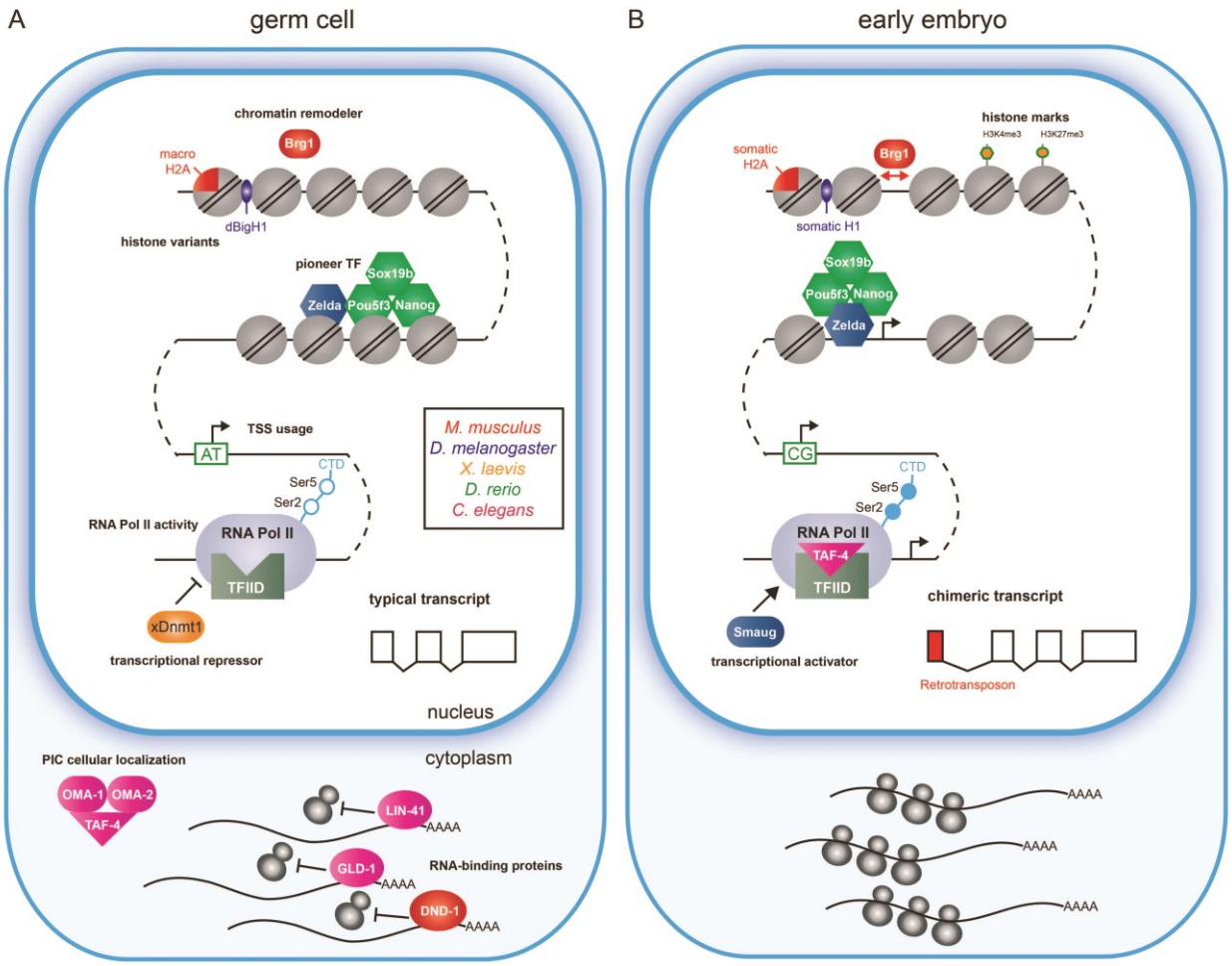


Figure 2.1 Integration of reprogramming-controlling mechanisms from different species in stereotypical cells

(A) germ cell: chromatin is in a closed conformation, mRNA transcription is inactive, mRNAs are translationally repressed. (B) early embryos: mRNA transcription of embryonic genes from open chromatin can occur, activating (H3K4me3) and repressive (H3K27me3) histone marks are deposited to regulate gene expression, mRNAs are translated into proteins. Colors of the involved factors match to the color of the respective model organism (indicated in the square box). References are listed in tables 1.3.1 and 1.3.2. Abbreviations: (TF) transcription factor, (TSS) transcription start site, (RNA Pol II) RNA Polymerase II, (PIC) pre-initiation complex, (CTD) C-terminal domain.

Type of regulation	Name	Organism	Function during OET	Molecular function	Reference
Histone variants	macroH2A	*M. musculus*	EGA repressor	Repressive H2A variant in germline and early embryo	(Chang et al. 2005)
	dBigH1	*D. melanogaster*	EGA repressor	Repressive H1 variant in germline and early embryo	(Perez-Montero et al. 2013)
Chromatin remodeler	Brg1	*M. musculus*	EGA activator	Catalytic subunit of the SWI/SNF chromatin remodeler complex	(Bultman et al. 2006)
Epigenetic modifier	xDnmt1	*X. laevis*	EGA repressor	DNA methyltransferase activity is not required for EGA repression, might act directly as transcriptional repressor	(Stancheva and Meehan 2000; Dunican et al. 2008)
Alternative promoters	unknown	*D. rerio*	both	Maternal A/T-rich and embryonic G/C-rich TSSs lead to differential activation of genes during OET	(Nepal et al. 2013; Haberle et al. 2014)
	Retro-transposons, especially LTRs	*M. musculus*	EGA activator	Retrotransposons act as alternative promoters in oocytes and embryos	(Peaston et al. 2004)
Basal transcription machinery	TAF-4	*C. elegans*	EGA activator	TFIID subunit that is repressed before EGA	(Guven-Ozkan et al. 2008)
TF with pioneer activity	Zelda	*D. melanogaster*	EGA activator	Activates transcription of embryonic genes, likely contains pioneer activity	(Liang et al. 2008; De Renzis et al. 2007; ten Bosch et al. 2006; Harrison et al. 2011; Nien et al. 2011)
	Nanog, Pou5f3, Sox19b	*D. rerio*	EGA activator	Activate transcription of embryonic genes, likely contain pioneer activity	(Lee et al. 2013; Leichsenring et al. 2013)

Table 2.1: Transcriptional regulation of reprogramming during the OET

2.1.5 Post-transcriptional regulation of reprogramming

RNA-binding proteins

As described earlier, germ cells have the developmental capacity to develop into cells of all three germ layers, which in disease, is manifested in teratomas. Posttranscriptional regulation appears to play a predominant role in preventing such a precocious onset of pluripotency in germ cells. In mouse, as in worms, post-transcriptional mechanisms involving the activity of RNA-binding proteins (RBPs) were found to maintain the germ cell fate and inhibit precocious reprogramming. In mouse, DND-1 (dead end 1) prevents teratoma formation in the embryonic germ cells (Youngren et al. 2005). Loss of DND-1 leads to a failure in downregulating pluripotency genes, like Nanog or Oct4, and defective cell cycle arrest (Cook et al. 2011). Defects in cell cycle arrest likely result from a failure to protect mRNAs encoding cell cycle regulators, like $p27^{Kip1}$ and $p21^{Cip1}$, from microRNA-mediated degradation (Kedde et al. 2007). Comparable to the function of DND-1 in mouse, two *C. elegans* RBPs named GLD-1 (defective in germline development 1) and LIN-41 (abnormal cell lineage 41) have been found to inhibit teratoma formation in the adult germ cells. Interestingly, GLD-1 and LIN-41 prevent germ cell reprogramming at two consecutive phases of oogenic differentiation. While GLD-1/Quaking functions earlier, maintaining the germ-line fate during the pachytene stage of meiosis I (Ciosk et al. 2006), LIN-41/TRIM-71 performs this function later, during the diplotene and diakinesis stages (Tocchini et al. 2014). Why different RBPs are employed at different developmental stages is not clear. One possibility is that LIN-41 takes the control from GLD-1 so that some of the GLD-1 mRNA targets, with OET-promoting functions, may be translated (Scheckel et al. 2012). Irrespective of the molecular details, common to *gld-1* and *lin-41* loss-of-function mutants is a premature entry of the meiotic cells into mitosis, followed by precocious EGA and teratomatous differentiation (Biedermann et al. 2009; Tocchini et al. 2014; Ciosk et al. 2006). In these mutants, the expression of early embryonic genes takes place before the expression of differentiation-specific genes (Tocchini et al. 2014). Moreover, just like in normal embryogenesis, the teratomatous differentiation, at least into muscles, depends on the maternally supplied transcription factor, PAL-1/Caudal, a master regulator for muscle fate in embryos (Baugh et al. 2005). All these observations suggest that the *gld-1* and *lin-41* teratomas reflect the function of GLD-1 and LIN-41 in delaying the onset of embryonic events until after fertilization. How precisely they do it remains unclear. One key GLD-1 target is the *cye-1* mRNA encoding cyclin E (Biedermann et al. 2009). By repressing this mRNA, and consequently preventing cyclin E translation, GLD-1 regulates the activity of a key cyclin E partner, the cell cycle-driving kinase CDK-2. In *gld-1* mutants, CDK-2 is critical for the precocious mitosis, which is somehow linked with EGA (Biedermann et al. 2009).

Similarly, in *lin-41* germlines, the translational de-repression of CDC-25.3, the activator of another cell cycle kinase called CDK-1, has been suggested to contribute to the premature entry into mitosis (Spike et al. 2014). Although in this case the connection between mitosis and EGA remains to be tested, GLD-1 and LIN-41 both repress premature mitosis, which appears to be linked to the transcriptional remodeling manifested by EGA and teratoma formation. Thus, a critical question for the future is whether there is a causal relation between the cell cycle and reprograming and, if yes, what is the underlying mechanism.

Although the most-studied posttranscriptional regulators act as the repressors of reprogramming, RBPs could also act as activators. In *Drosophila*, another RBP, Smaug, has been reported to act as an EGA activator. The loss of Smaug leads to pleiotropic developmental defects in embryos, including the failure to start high-level embryonic transcription, possibly due to defects in the activation of the basal transcription machinery (Benoit et al. 2009). However, whether Smaug directly regulates the transcription machinery, or whether the failure to activate transcription is indirectly due to additional functions of Smaug in maternal mRNA degradation (Tadros et al. 2007), remains to be seen.

P granules

C. elegans P granules are germline-specific ribonucleoprotein granules with, typically, perinuclear localization, similar to germ granules or nuage in other organisms. It has been observed that, in *gld-1* and *lin-41* mutant germlines, but also in germlines undergoing a direct germ-to-soma-conversion (Tursun et al. 2011), the P granules disappear from differentiating cells. However, whether the loss of P granules is the cause or consequence of the somatic differentiation of germ cells remains controversial. A recent study, aimed at answering this question, used simultaneous knock-down of RBPs that nucleate P granule formation (Updike et al. 2014). This resulted in the expression of some somatic genes and differentiation into neurons after the ectopic expression of a neural fate master regulator called CHE-1. Thus, it seems that compromising P granules enhances the ability of germ cells to be reprogrammed into somatic cells, but, alone, is not sufficient to cause an OET-like transition, as observed in *gld-1* and *lin-41* mutant germ lines. Also, while in some cases the association of a given RBP with P granules may be important for the RBP activity (Voronina et al. 2012), in other cases, including GLD-1 or LIN-41, the significance of the association with P granules (if any) remains unknown. Finally, compromising P granules could affect the biogenesis of endo-siRNAs (endogenous small interfering RNAs), including those facilitating the expression of germ-line genes (Gu et al. 2009; Cecere et al. 2014; Claycomb et al. 2009). Thus, the connection between P granules and the maintenance of the germ cell fate may be indirect.

In summary, post-transcriptional regulation appears to contribute to both, EGA/pluripotency repression in germ cells and EGA activation in early embryos (the involved factors are illustrated in Figure 2.1 and summarized in Table 2.2). RBPs have been found to serve as cytoplasmic roadblocks to reprogramming and thereby prevent a precocious pluripotent state in germ cells. Their identification in mice and worms suggests that this is likely a conserved regulatory paradigm. This is further supported by the common use of post-transcriptional regulation in oocytes, as oocyte development is associated with the entry into a transcriptionally silent state in most species. Common targets regulated by these RBPs are cell cycle genes, which indicates that they could be conserved targets for reprogramming regulation. The role of RNP granules in germ cell identity maintenance remains less well understood, as their impairment likely disturbs the function of many cytoplasmic regulators.

Type of regulation	Name	Organism	Function during OET	Molecular function	Reference
RBPs	DND-1	*M. musculus*	EGA repressor	Translational activator of the cell-cycle genes $p27^{Kip1}$ and $p21^{Cip1}$, translational repressor of pluripotency genes like Nanog	(Youngren et al. 2005; Cook et al. 2011)
	GLD-1	*C. elegans*	EGA repressor	Translational repressor of the cell-cycle gene CYE-1/CyclinE	(Ciosk et al. 2006; Biedermann et al. 2009)
	LIN-41	*C. elegans*	EGA repressor	Translational repressor	(Tocchini et al. 2014)
	Smaug	*D. melanogaster*	EGA activator	Activator of the basal transcription machinery, maternal mRNA degradation	(Benoit et al. 2009; Tadros et al. 2007)
Germ granules	PGL-1, PGL-3, GLH-1, GLH-4	*C. elegans*	Somatic gene repressor	Restrict the reprogramming of germ cells into somatic cells	(Updike et al. 2014)

Table 2.2: Post-transcriptional regulation of reprogramming during the OET

2.2 *Caenorhabditis elegans*: an elegant model to study the oocyte-to-embryo transition

Scientists have been using the nematode *Caenorhabditis elegans* to study many different aspects of animal development. Its short life cycle of approximately three days, the easy cultivation, the fully sequenced genome and a variety of genome editing tools provide many advantages. Features that make *C. elegans* especially suited to study germ cell development and the oocyte-to-embryo transition (OET) are its transparency, which allows direct observation of developing germ cells and embryos in live worms, the large number of progeny of up to 300 per hermaphrodite and the linearly ordered gonad with consecutive developmental stages. The *C. elegans* reproductive system, and in particular the process of oocyte maturation and the mechanisms controlling transcription during the OET are described below.

2.2.1 The reproductive system

C. elegans can have two different sexes, hermaphrodites and males. Hermaphrodites produce sperm during their last larval stage, which are then stored in the spermatheca. Adult hermaphrodites only form oocytes, which get self-fertilized by their stored sperm, unless they mate to males. Male sperm have specific features with which they can outcompete the hermaphrodite sperm, thereby enabling the exchange of genetic material between animals. The reproductive system of the hermaphrodite consists of two symmetrical U-shaped gonad arms ending in a common uterus (Figure 2.2). The distal part of the gonad contains actively dividing nuclei in a syncytium and provides a continuous supply of cells to create the gametes (Ellis and Kimble 1994). This pool of stem cells is controlled by the distal tip cell (DTC), which surrounds the distal gonad and sends proliferation signals (Kimble and White 1981). In the medial gonad region, the cells enter meiosis and undergo cellularization around the loop region. These cells differentiate into oocytes and arrest at diakinesis of meiotic prophase. The very proximal oocyte before the spermatheca (also called -1 oocyte) undergoes maturation and is finally pushed into the spermatheca to get fertilized. This process is very rapid with one oocyte getting fertilized every 23 minutes (McCarter et al. 1999). The zygote is then released into the uterus, where it completes several embryonic divisions until it is laid by the hermaphrodite.

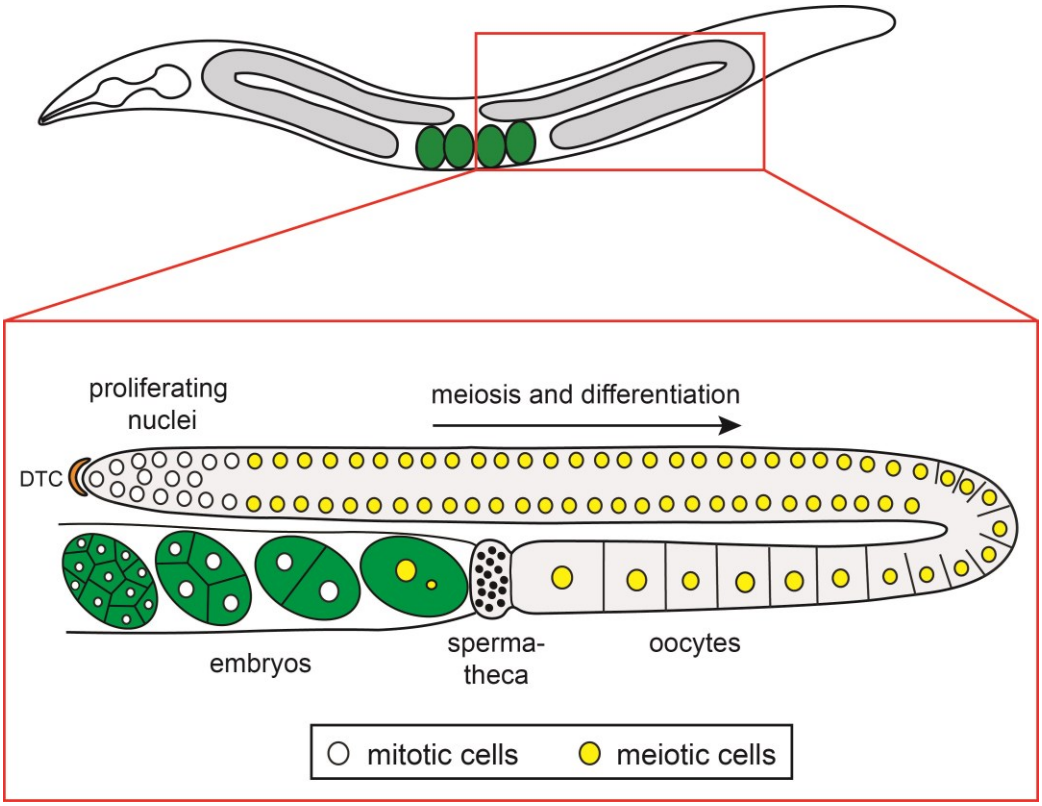


Figure 2.2: The *Caenorhabditis elegans* reproductive system

C. elegans hermaphrodites have two U-shaped gonad arms that end in a common uterus, where the embryos undergo early development, before being laid. The red box shows a close-up of one gonad arm. The distal tip cell (DTC) promotes proliferation of the germline stem cells. Subsequently, the mitotic cells enter meiosis and undergo differentiation into oocytes in the adult hermaphrodite. One oocyte after the other gets fertilized in the spermatheca and gives rise to a zygote that completes the meiotic division and then starts to undergo rapid mitotic divisions.

2.2.2 Oocyte maturation

The production of intact mature oocytes is required for a successful oocyte-to-embryo transition. Many coordinated processes drive the meiotic maturation of oocytes, which is defined as the cell cycle progression from diakinesis to metaphase of meiosis I. Cellular processes occurring during the meiotic maturation include nuclear envelope breakdown (NEBD), rearrangement of the cytoskeleton, and meiotic spindle assembly (McCarter et al. 1999). These processes are initiated by activation of the Maturation Promoting Factor (MPF), which is a well conserved key regulator of the meiotic cell cycle progression

(Masui 2001). The mechanisms regulating MPF activity are summarized below based on the knowledge from vertebrate systems, however, these mechanisms appear to be well conserved also in *C. elegans* (Von Stetina and Orr-Weaver 2011). MPF is composed of the Cyclin-dependent kinase Cdk1 and its regulatory subunit Cyclin B and is kept inactive in arrested oocytes by inhibitory phosphorylations at Thr14 and Tyr15 on Cdk1 (Kim et al. 2013) (Figure 2.3). These inhibitory phosphorylations are added by the Wee1/Myt1 kinases (Mueller et al. 1995) and can be removed by the phosphatase Cdc25 (Kumagai and Dunphy 1991) upon hormonal stimulation. Removal of the inhibitory phosphorylations activates MPF, which can then phosphorylate downstream substrates regulating the key cellular processes during meiotic maturation as described above. Metaphase exit requires the function of the anaphase-promoting complex (APC), an E3 ubiquitin ligase, which induces cyclin B degradation and thereby leads to inactivation of MPF (Kim et al. 2013).

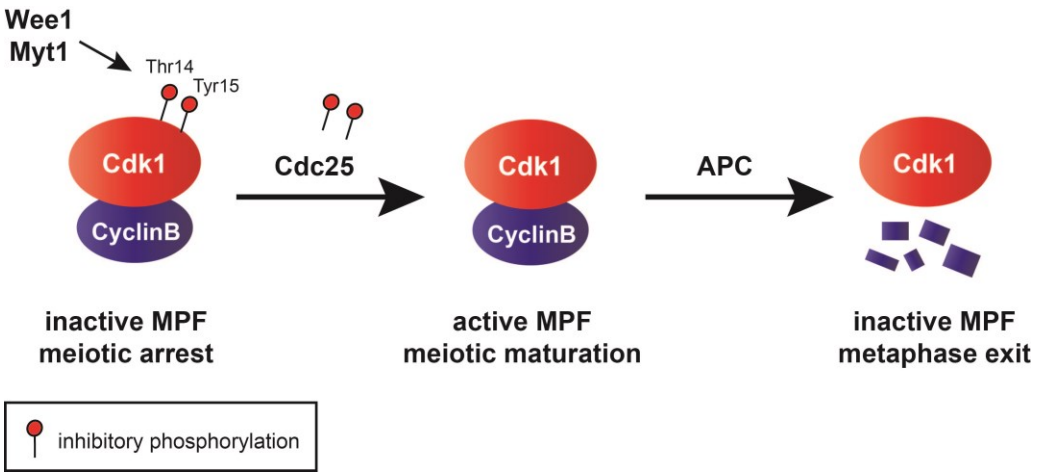


Figure 2.3: Regulation of the Maturation Promoting Factor in meiotic oocytes

MPF is the key regulator of meiotic cell cycle progression. Its activity is regulated through phosphorylation, dephosphorylation and cyclin degradation. Abbreviations: (MPF) Maturation Promoting Factor, (Cdk1) Cyclin-dependent kinase 1, (APC) Anaphase-Promoting Complex.

In *C. elegans,* meiotic maturation is stimulated by the major sperm protein (MSP) (Miller et al. 2001; McCarter et al. 1999). Other important functions of MSP signaling in reproduction include the induction of gonadal sheath cell contraction that promotes ovulation and the activation of the mitogen-activated protein kinase (MAPK) signaling pathway (Miller et al. 2001). In addition, the RNA-binding proteins OMA-1/-2 appear to be important for meiotic maturation and ovulation (Detwiler et al. 2001; Shimada et al. 2002). Loss of OMA-1/-2 proteins leads to MAPK inactivation, defects in NEBD and a failure in the

localization of AIR-2, the *C. elegans* Aurora B kinase, to chromatin. OMA proteins appear to function upstream of the MPF in meiotic maturation (Detwiler et al. 2001).

2.2.3 Transcriptional control during the oocyte-to-embryo transition

A global repression of mRNA transcription in oocytes is common among diverse species. In *C. elegans* mRNA transcription is shut down when oocytes enter diakinesis and the transcriptional quiescence is maintained until the four-cell stage embryo (Kelly et al. 2002; Schisa et al. 2001; Seydoux and Dunn 1997). In embryonic germ cell precursors, transcription continues to be silenced until after the 100-cell stage (Mello et al. 1992). The event of transcription initiation in the early embryo is called embryonic genome activation (EGA). The transcriptional and post-transcriptional mechanisms controlling EGA in different animal models have been summarized in the previous chapter. Here, I focus just on those mechanisms that have been identified to regulate the general transcription machinery during the OET in *C. elegans* (summarized in Figure 2.4). Polymerase II (Pol II)-mediated mRNA transcription is a multi-step process that starts with the pre-initiation complex (PIC) assembly, consisting of Pol II, general transcription factors (GTFs) and the multi-protein complex Mediator (Blackwell and Walker 2006). Subsequent transcription steps depend on the phosphorylation of the Pol II C-terminal domain (CTD) on serine 5 (PSer5) during initiation and serine 2 (PSer2) during elongation, and are followed by transcription termination. *C. elegans* germ cells are transcriptionally active until diakinesis, when transcription is shut down by an unknown mechanism. Meiotic maturation, occurring in the last oocyte before the spermatheca, then appears to stimulate transcription initiation specifically. This has been observed by the accumulation of PSer5, but not PSer2, in *wee-1.3*-depleted oocytes, which undergo precocious maturation, due to premature CDK-1 activation (Walker et al. 2007). Maturation-dependent PSer5 accumulation in oocytes is inhibited by FCP-1, a phosphatase that likely recycles phosphorylated Pol II, and by a ubiquitination pathway involving the E1 ligase UBA-1 (Walker et al. 2007). Transcriptional inhibition in the one-and two-cell embryo is mediated via inhibition of TAF-4, a subunit of the general TFIID, by sequestration of TAF-4 in the cytoplasm by OMA-1/-2 zinc finger proteins (Guven-Ozkan et al. 2008). Degradation of the OMA proteins in the four-cell embryo leads to release of TAF-4 into the nucleus, where transcription can start. Transcription repression in the embryonic germ cells precursors is maintained by PIE-1, another zinc finger protein (Mello et al. 1992; Seydoux and Dunn 1997). PIE-1 interferes with transcription elongation by inhibiting CDK-9, the kinase responsible for Ser2 phosphorylation in embryos (Zhang et al. 2003). Consistent with a block in elongation, Pser5 positive loci are present in those germ cell precursors (Seydoux and Dunn 1997).

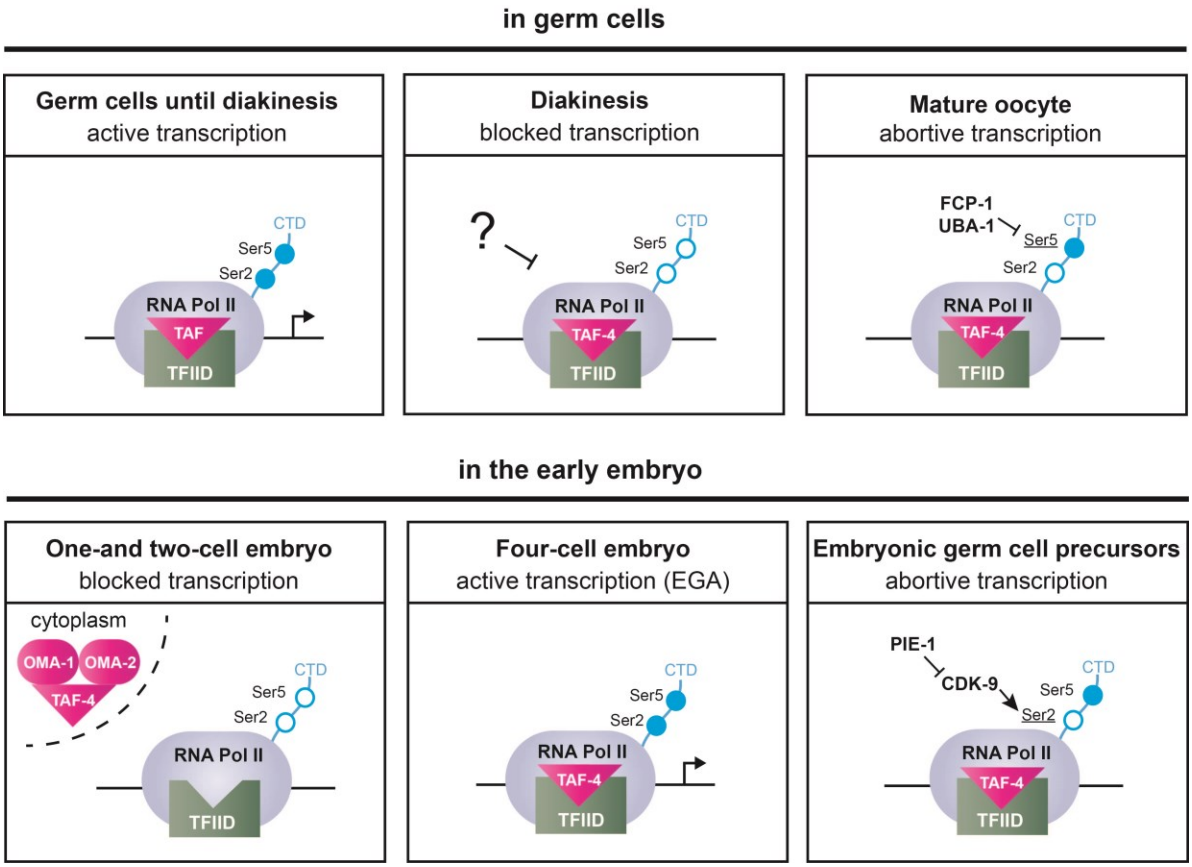


Figure 2.4: Transcriptional control during the OET in *C. elegans*

Germ cells are transcriptionally active until diakinesis, when transcription is inhibited by an unknown mechanism (filled dots indicate phosphorylated serines, empty dots indicate unphosphorylated serines). Maturation induces transcription initiation, indicated by Ser5 phosphorylation, but not elongation. Accumulation of Pser5 in mature oocytes is inhibited by the phosphatase FCP-1 and an ubiquitin pathway involving the E1 ligase UBA-1. In one-and two-cell embryos, transcription is inhibited by sequestration of TAF-4, a TFIID subunit, in the cytoplasm. EGA is initiated in the four-cell embryo, except for the embryonic germ cell precursor, where transcriptional silencing is maintained by PIE-1. The upper row summarizes the transcriptional states in germ cells, the lower row in embryonic blastomeres. Abbreviations: (RNA Pol II) RNA Polymerase II, (CTD) C-terminal domain.

2.3 Small RNA pathways

Small RNAs have been proven to be of fundamental importance for many biological processes across model organisms. They serve as an adaptive immune system against viral infections, regulate development and life span, prevent transposon mobilization and modulate the expression of endogenous genes, among others. In eukaryotes there are three major classes of small RNAs: microRNAs (miRNAs),

Piwi-interacting RNAs (piRNAs) and small interfering RNAs (siRNAs). Although these classes differ in their biogenesis and their mode of target regulation, they share mechanistic principles in which small RNAs guide Argonaute proteins to complementary target sequences to regulate their expression at the transcriptional or post-transcriptional level. The induction of gene silencing by antisense RNA, a mechanism known as RNA interference (RNAi), was first discovered in *Caenorhabditis elegans* (Fire et al. 1998). Since then, the nematode proved to be a very valuable model organism to study small RNA biology. Recent advances include especially the discovery of the diversity of endogenous siRNAs and their regulatory mechanisms. Endogenous siRNAs can be distinguished from exogenous siRNAs by the source of RNA they are made of. Exogenous siRNAs are processed from exogenous dsRNA and mainly play a role in the antiviral defense mechanism. Endogenous siRNAs are produced from endogenous dsRNA and regulate the expression of repetitive elements, pseudogenes and mRNAs. Endogenous siRNAs predominantly appear to play a role in the germline. In the *C. elegans* germline two siRNA pathways have been identified, the WAGO (*worm argonaute*) pathway and the CSR-1 (*chromosome-segregation and RNAi-deficient-1*) pathway. Although the two pathways share biogenesis components (Table 2.3), they differ considerably in their targets and functional outputs and will be therefore described separately in the following sections.

2.3.1 The WAGO pathway

The WAGO pathway is named after the worm-specific Argonaute proteins, which load the endo-siRNAs in this pathway. It mostly silences its targets, including repetitive sequences, transposons and a few protein-coding mRNAs (Gu et al. 2009). Silencing of WAGO targets occurs at the post-transcriptional and transcriptional level and can even result in RNA-induced epigenetic silencing (RNAe), the multigenerational inheritance of silencing (Buckley et al. 2012). The WAGO pathway can be further separated into three different upstream pathways that commonly use the WAGOs as downstream effectors (Figure 2.5). Two of these upstream pathways are dependent on ERI-1 (*enhanced RNAi*), a 3'-to-5'-exonuclease (Kennedy et al. 2004; Lee et al. 2006). ERI-1 interacts with DCR-1 (*dicer related-1*) and forms together with the RNA-dependent RNA polymerase (RdRP) RRF-3, and additional components, the ERI (ERI-DCR-1) complex (Thivierge et al. 2011). The ERI complex produces a set of endogenous small RNAs that are 26 nucleotides in length with a guanosine at the 5'end (26G RNAs) (Ruby et al. 2006; Han et al. 2009). Two different populations of 26G RNAs exist: sperm-26G RNAs, which are bound by the redundant Argonautes ALG-3 and ALG-4, and oocyte-26G RNAs, which are bound by ERGO-1 (Han et al. 2009; Conine et al. 2010; Vasale et al. 2010). ALG-3/4 26G RNAs are required for spermiogenesis, and loss of ALG-3/4

results in a temperature-dependent sterility of the animals (Han et al. 2009; Conine et al. 2010). ERGO-1 26G RNAs appear to target predominantly zygotic transcripts during development and map to gene-poor chromosome arms and duplicated non-conserved genes (Han et al. 2009; Vasale et al. 2010; Fischer et al. 2011). Loss of ERGO-1 only leads to minor defects in fertility, but mutant worms show the ERI phenotype (Yigit et al. 2006; Han et al. 2009). The third upstream pathway, which uses WAGOs as secondary effectors, is the PIWI-interacting RNA (piRNA) pathway. piRNAs are germ-line specific small RNAs of 21 nucleotides length with a 5' uridine (Ruby et al. 2006). Their biogenesis differs considerably from that of 26G RNAs, in that they are produced by RNA Pol II. In addition, whereas 26G RNAs show perfect complementarity to their target transcripts, piRNAs only have a partial complementarity (Bagijn et al. 2012). Their production is initiated by Forkhead transcription factors (FKH) from two loci on chromosome IV or from promoter regions of protein-coding genes genome-wide (Ruby et al. 2006; Cecere et al. 2012). piRNAs associate with the PIWI protein PRG-1, whose loss results in a temperature-dependent sterility (Batista et al. 2008). Due to their role in transgene silencing, piRNAs are thought to silence "non-self sequences" and to participate in genome surveillance (Bagijn et al. 2012; Lee et al. 2012). However, while piRNAs can initiate the silencing, WAGO 22G RNAs are needed to maintain the silencing response (Lee et al. 2012; Ashe et al. 2012; Shirayama et al. 2012). WAGO 22G RNAs are named after their length of 22 nucleotides and the presence of a guanosine at the 5'end and they are loaded by 12 semiredundantly functioning Argonautes of the WAGO clade (Gu et al. 2009). Loss of all 12 Argonautes in the MAGO 12 mutant leads to a depletion of basically all WAGO 22G RNAs and a temperature-dependent sterility of the animal (Gu et al. 2009).The production of WAGO 22G RNAs is DCR-1-independent, and occurs via RNA-dependent RNA polymerase (RdRP) complexes. These consist of the redundantly functioning RdRPs RRF-1 and EGO-1, the dicer-related helicase DRH-3, and a tudor-domain containing protein called EKL-1 (Gu et al. 2009). In addition, the WAGO 22G RNA production is partially dependent on the so-called Mutator foci, which are formed by the 6 Mutator proteins MUT-16, MUT-7, MUT-8/RDE-2, MUT-2/RDE-3, MUT-15, and MUT-14 (Billi et al. 2014). Loss of any of these mutator proteins results in transposon upregulation, leading to a high germline mutation rate and sterility after several generations. Mutator foci likely serve as amplification centers for WAGO 22G RNAs by recruiting RdRP components (Billi et al. 2014). The WAGO pathway can induce downstream signaling *via* the nuclear RNAi pathway. This pathway mediates the transcriptional silencing of targets in the soma and germ line, via WAGO-12/NRDE-3 and WAGO-9/HRDE-1, respectively (Ashe et al. 2012; Buckley et al. 2012; Guang et al. 2010).

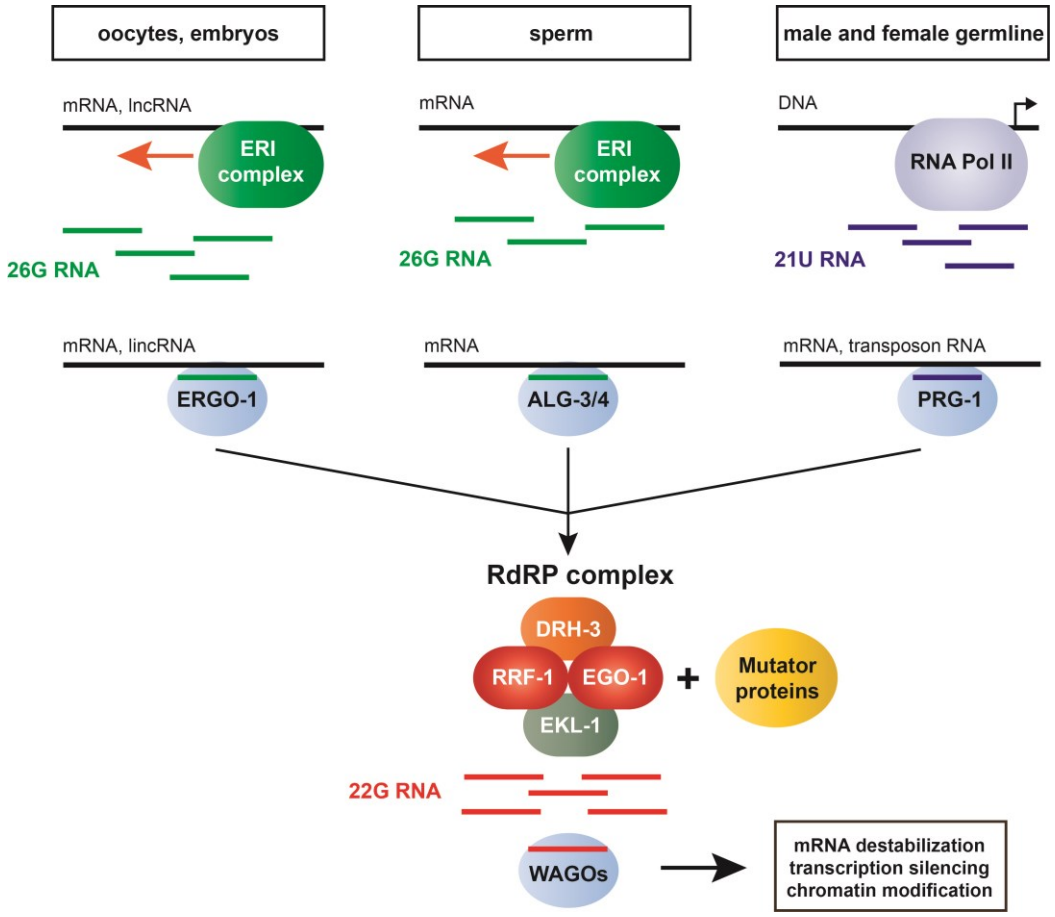


Figure 2.5: The WAGO pathway

Different primary small RNA pathways function upstream of the WAGO pathway. They include the ERI-1-dependent ERGO-1 and ALG-3/4 26G RNA pathways and the 21U-piRNA pathway. Biogenesis of WAGO 22G RNAs occurs via the RdRPs RRF-1 and EGO-1 and is dependent on the Mutator proteins. 22G RNAs are bound by WAGOs to associate with the targets of the respective upstream pathways. Abbreviations: (mRNA) messenger RNA, (lncRNA) long non-coding RNA, (RNA Pol II) RNA Polymerase II, (RdRP) RNA-dependent RNA polymerase, (WAGO) worm-specific Argonaute.

Small RNA	5'terminus	Argonaute	Tissue	Polymerase	Biogenesis factors
ERGO-1 26G RNA	monophosphate	ERGO-1	germline (oocyte), soma (embryo)	RRF-3	DRH-3, ERI-5, DCR-1, ERI-1b, ERI-3, RDE-4
ALG-3/4 26G RNA	monophosphate	ALG-3/4	germline (sperm)	RRF-3	DRH-3, ERI-5, DCR-1, ERI-1b, ERI-3, RDE-4
WAGO 22G RNA	triphosphate	WAGOs	germline, soma	RRF-1, EGO-1	DRH-3, EKL-1, Mutator proteins
CSR-1 22G RNA	triphosphate	CSR-1	germline	EGO-1	DRH-3, EKL-1
21U piRNA	monophosphate	PRG-1/2	germline	RNA Pol II	FKH-3/4/5, UNC-130

Table 2.3: Features of the different endogenous siRNA pathways in *C. elegans*

2.3.2 The CSR-1 pathway

Although CSR-1 can induce target mRNA degradation via its slicing activity (Gerson-Gurwitz et al. 2016), the CSR-1 pathway is suggested to not commonly silence its targets, but rather promote their expression on the transcriptional level (Cecere et al. 2014). Consistent with this positive role on gene expression, the CSR-1 pathway is predicted to counteract the gene silencing mediated by the piRNA pathway by recognizing and licensing self-sequences for expression (Conine et al. 2010; Seth et al. 2013; Wedeles et al. 2013b). The current model for such gene activation is that CSR-1 interacts with Polymerase II in an RNA-dependent manner and influences chromatin organization, possibly by recruiting chromatin modifiers and histone methyltransferases that promote euchromatin formation (Wedeles et al. 2013a). Though, the precise mechanism of the gene activation remains unclear and chromatin modifiers interacting with CSR-1 have not been identified to date. The majority of CSR-1 22G RNAs are antisense to protein-coding genes, more specifically to thousands of germline genes (Claycomb et al. 2009). What

initially triggers their production is unclear, as no primary siRNAs or Argonautes in this pathway have been identified yet. Like the WAGO pathway, the CSR-1 pathway uses an RdRP complex for the biogenesis of 22G RNAs (Figure 2.6). However, whereas WAGO 22G RNAs are redundantly produced by the RdRPs RRF-1 and EGO-1, CSR-1 22G RNAs depend on EGO-1 alone for their production (Gu et al. 2009; Claycomb et al. 2009). DRH-3 and EKL-1 are shared components of the RdRP complexes for the WAGO and CSR-1 biogenesis. The 22G RNAs are loaded into CSR-1, the only Argonaute functioning in this pathway (Claycomb et al. 2009). The abundance of a subset of CSR-1 22G RNAs is controlled by CDE-1, which uridylates and thereby destabilizes the 22G RNAs (van Wolfswinkel et al. 2009). CDE-1 is suggested to have a role in the separation of the WAGO and CSR-1 pathway, as the accumulation of CSR-1 22G RNAs in *cde-1* mutants leads to defects in gene silencing, likely due to misloading of CSR-1 22G RNAs into WAGOs (van Wolfswinkel et al. 2009). Loss of CSR-1 is reported to result only in modest changes in gene expression (Claycomb et al. 2009), which is in contrast to the severe developmental defects that lead to sterility and embryonic lethality. These phenotypes are shared among mutants of the CSR-1 pathway, such as DRH-3 and EGO-1. Germline defects specifically include underproliferation of mitotic cells, enlargement of the transition zone from mitotic to meiotic cells, chromosome segregation defects in diakinetic oocyte nuclei, and impaired sperm development (Smardon et al. 2000; Vought et al. 2005; Claycomb et al. 2009; Qiao et al. 1995; Conine et al. 2010). Embryonic phenotypes include poor metaphase alignment and chromosomal bridging at anaphase resulting in abnormally shaped nuclei and embryonic arrest before gastrulation (Smardon et al. 2000; Claycomb et al. 2009). These meiotic and mitotic defects have been originally attributed to the failures in chromatin organization. A significant contribution to embryonic lethality is suggested to come from the depletion of histones in CSR-1 pathway mutants, as overexpression of the core histones improved the survival rate of *csr-1*- and *ego-1*-depleted embryos (Avgousti et al. 2012). However, the recently reported CSR-1 endonuclease (slicing) activity, also contributes to normal embryogenesis, presumably by fine-tuning maternal mRNAs encoding for proteins with essential functions in the embryo (Gerson-Gurwitz et al. 2016). It has been observed, that increased levels of a microtubule depolymerase, called KLP-7, in CSR-1 slicer mutants lead to the microtubule defects observed in *csr-1*-depleted embryos. Required for the slicing activity is the DDH catalytic motif, which has been also shown to mediate mRNA cleavage by CSR-1 *in vitro* (Aoki et al. 2007). An additional negative regulatory role for CSR-1, is the translational repression of mRNAs (Friend et al. 2012). CSR-1 has been shown to interact with the PUF protein FBF-1 and the translation elongation factor eEF1A to repress FBF-1 target mRNAs. It remains unclear if this function is small RNA-dependent as depletion of other CSR-1 pathway components resulted in a less robust de-repression of FBF-1 targets.

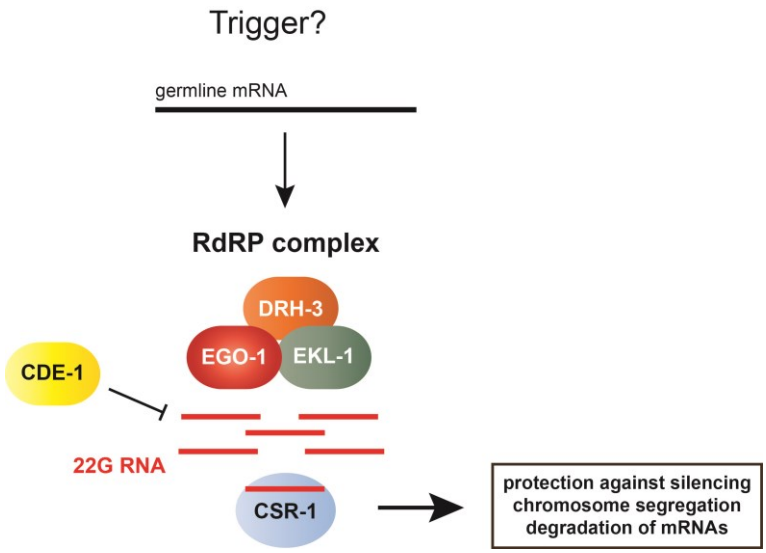


Figure 2.6: The CSR-1 pathway

It is not known what initially triggers the production of CSR-1 22G RNAs by the RdRP EGO-1. CSR-1 22G RNAs are antisense to more than 4000 germline-expressed genes. CSR-1 is the only Argonaute protein functioning in this pathway, which appears to usually not silence its targets, but rather promotes the expression of the germline-expressed genes. However, also negative regulatory functions of CSR-1 have been found, including the degradation of target mRNAs *via* its slicing activity. CDE-1 regulates the abundance of CSR-1 22G RNAs by triggering their degradation *via* uridylation. Abbreviations: (mRNA) messenger RNA, (RdRP) RNA-dependent RNA polymerase.

3 Results and Discussion

3.1 Inhibition of embryonic gene expression in the *Caenorhabditis elegans* germline by the CSR-1 small RNA pathway

with Cristina Tocchini and Michael B. Stadler

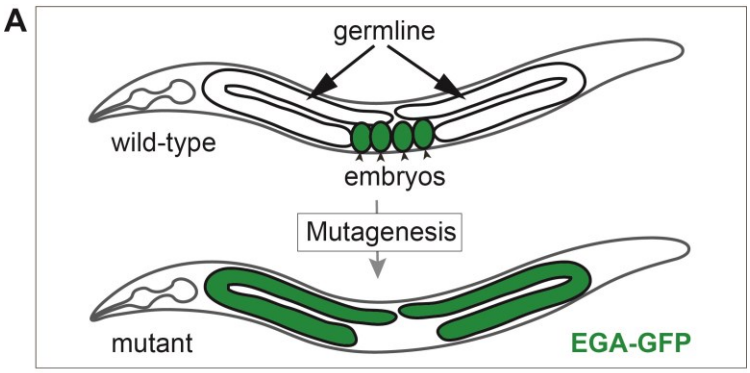
In order to give rise to a new organism, the blastomeres of the early embryo must be developmentally flexible. This developmental plasticity, or pluripotency, is acquired during the so-called oocyte-to-embryo transition (OET), when a terminally differentiated cell, the oocyte, is reprogrammed, after fertilization, into an embryo constituted by undifferentiated and pluripotent cells. A precise regulation of this reprogramming event is critical, since germ cells that precociously acquire pluripotency can develop into germ cell tumors, called teratomas (Ciosk et al. 2006). This disease phenotype has been observed across species, from the nematode *Caenorhabditis elegans* to humans. In *C. elegans,* a teratoma is induced in the absence of the RNA binding proteins (RBPs) GLD-1/Quaking (Ciosk et al. 2006) or LIN-41/TRIM71 (Tocchini et al. 2014). GLD-1 and LIN-41 inhibit a precocious germ cell reprogramming at consecutive meiotic stages of oocyte development, in pachytene-stage and diplotene/diakinesis-stages, respectively. Similar to the naturally occurring events during the OET, the reprogramming in *gld-1* and *lin-41* mutant germ cells is characterized by a re-entry into the cell cycle, followed by embryonic genome activation (EGA), and finally somatic-like differentiation into various cell types, such as muscles and neurons (Tocchini et al. 2014; Biedermann et al. 2009; Ciosk et al. 2006). Although the identification of GLD-1 and LIN-41 contributed to elucidating the mechanisms controlling pluripotency during the OET, our understanding is still far from being complete. Therefore, we decided to screen for additional repressors of pluripotency in the *C. elegans* germline. Interestingly, we find mutants showing EGA in the germline (later referred to as germline-EGA) without a cell cycle re-entry. In addition, the germlines of these mutants show expression only of early embryonic genes, but not of later-expressed differentiation-promoting genes. The mutant alleles were mapped to the *drh-3* and *ego-1* genes, suggesting an involvement of endogenous small interfering RNA (endo-siRNA) pathways in embryonic gene repression in germ cells. By testing multiple endo-siRNA mutants for germline-EGA, we find that in particular one of the two *C. elegans* endo-siRNA pathways, employing the Argonaute CSR-1, is important for such a function. CSR-1 is the only worm Argonaute required for fertility and embryo survival (Yigit et al. 2006). It

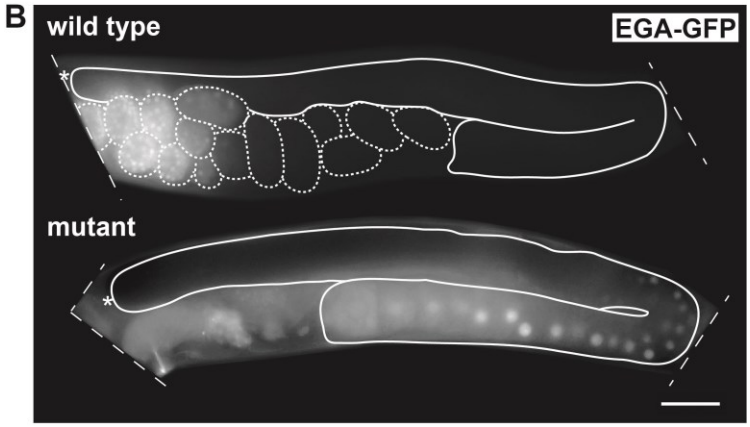
has been suggested to function in diverse processes, including chromosome segregation (Claycomb et al. 2009; Yigit et al. 2006), chromatin organization (Maine et al. 2005; She et al. 2009), histone mRNA processing (Avgousti et al. 2012), germ granule formation (Claycomb et al. 2009; Updike and Strome 2009), alternative splicing (Barberan-Soler et al. 2014), and exogenous RNAi (Yigit et al. 2006). CSR-1 binds small RNAs that are complementary to germline-expressed genes (Claycomb et al. 2009) and it is suggested to promote expression of its target genes (Cecere et al. 2014). Moreover, CSR-1 is predicted to counteract gene silencing, by recognizing and licensing self-sequences for expression (Seth et al. 2013; Wedeles et al. 2013b; Conine et al. 2013). Reported negative regulatory roles of CSR-1 comprise the translational repression of FBF-1 target mRNAs in mitotic germ cells (Friend et al. 2012), and the degradation of maternal mRNAs *via* its conserved catalytic slicer residues (Gerson-Gurwitz et al. 2016; Yigit et al. 2006). In this study, we present a new role for CSR-1 in inhibiting the precocious expression of embryonic genes in the *C. elegans* germline, and show that this function is mediated *via* its slicing activity. These findings suggest that CSR-1 is important during the OET to delay embryonic transcription to its physiological onset in early embryos.

3.1.1 Results

Identification of new pluripotency repressors

To better understand how pluripotency is controlled in the *C. elegans* germline we performed a genetic screen to discover new pluripotency repressors. It was previously observed that the reprogramming of germ cells to pluripotency in *gld-1* mutants was accompanied by a precocious onset of EGA (Biedermann et al. 2009), linking pluripotency to EGA. Moreover, the identification of pluripotency repressors based on a precocious EGA in germ cells has been recently proven to be a valuable tool, through the identification of the RNA binding protein LIN-41 as a pluripotency repressor (Tocchini et al. 2014). Therefore, in this screen, we continued to use EGA as a marker for pluripotency. EGA is monitored with an EGA-GFP reporter, expressing GFP from the promoter of a very early expressed embryonic gene (*vet-4*), resulting in an embryo-specific expression in wild type worms. To identify novel regulators, we screened for mis-expression of the EGA reporter in the germline (Figure 3.1A). We discovered three independent mutants (alleles *rrr2*, *rrr5*, and *rrr9*) expressing the EGA-GFP reporter precociously in germ cells (Figure 3.1B). By performing complementation group assays between our newly identified mutants, as well as with *gld-1* and *lin-41* mutants, we verified the novelty of the identified regulators, and discovered that the EGA phenotype in the *rrr2* and the *rrr5* mutant is caused by an alteration in the same gene (Figure 3.1C).





C

Allele	*gld-1 (q485)*	*lin-41 (rrr3)*	*rrr2*	*rrr5*	*rrr9*
rrr9	WT	WT	WT	WT	**EGA**
rrr5	WT	WT	**EGA**	**EGA**	WT
rrr2	WT	WT	**EGA**	**EGA**	WT
lin-41 (rrr3)	WT	**EGA**	WT	WT	WT
gld-1 (q485)	**EGA**	WT	WT	WT	WT

Figure 3.1: Identification of new pluripotency repressors

A. Schematic representation of the genetic screen enrolled to identify pluripotency repressors in the adult *C. elegans* germline. As a readout for pluripotency an EGA-GFP reporter was used (green, showing embryonic transcription). Wild-type animals express the GFP reporter in embryos. Mutant animals abnormally express the reporter in germ cells. **B.** Fluorescent micrographs of live animals expressing the EGA-GFP reporter. Here, and in the subsequent figures, gonads are outlined with a continuous line, embryos with a dashed line and asterisks mark the distal end of the gonad. Wild type animals express the embryonic reporter specifically in embryos, whereas mutant animals precociously express it in developing oocytes. Scale bar: 40 µm. **C.** Complementation matrix of newly identified mutants *rrr2/rrr5/rrr9* and previously identified pluripotency repressors *gld-1* and *lin-41*. For the complementation crosses we used the putative null alleles *gld-1(q485)* and *lin-41(rrr3)*. "WT" indicates complementation, "EGA" indicates non-complementation. The mutant alleles *rrr2* and *rrr5* do not complement each other.

EGA occurs independently of the cell-cycle

Previously, EGA was shown to occur in *gld-1* and *lin-41* mutant germ cells that start to undergo meiosis, but then re-enter into the mitotic cell cycle and differentiate into a teratoma (Francis et al. 1995; Ciosk et al. 2006; Tocchini et al. 2014). In contrast to the *gld-1* and *lin-41* phenotypes, the newly identified mutants were able to give rise to oocyte-like cells with condensed chromosomes (Figure 3.2A), suggesting the entry of the cells into diakinesis and a normal progression through meiosis. In wild-type oocytes, centrosomes are eliminated during meiotic progression (Mikeladze-Dvali et al. 2012). If the mutant oocytes undergo a normal meiosis, we thus expected the elimination of centrosomes in those cells. We monitored the presence of centrosomes by staining for the centrosome component SPD-2. We observed that SPD-2 was absent in *rrr2* mutant oocytes, like it was in wild-type oocytes (Figure 3.2A), indicating the normal elimination of centrosomes and a meiotic progression of the mutant oocytes. In addition to being a subsequent event of the re-entry of germ cells into mitosis, EGA has also been observed as a result of precocious oocyte maturation in *wee-1.3*-depleted oocytes (Allen et al. 2014). Oocyte maturation, the exit of meiosis in oocytes, occurs in wild-type gonads in the last maturing oocyte, and is characterized by nuclear envelope breakdown (NEBD), cytoskeletal rearrangements and meiotic spindle assembly (McCarter et al. 1999). To test, if EGA in the mutant oocytes is the result of a precocious oocyte maturation, we examined their nuclear envelopes. To visualize nuclear envelopes, we followed the localization of EMR-1, a protein present in nuclear envelopes, by using an *emr-1-mCherry* transgene (Morales-Martinez et al. 2015). NEBD was observed, as reported before, in maturing oocytes in wild type animals (Figure 3.2B). Mutant oocyte-like cells expressing EGA-GFP, however, showed intact nuclear envelopes (Figure 3.2B). In sum, the mutant oocytes appeared to progress normally through meiosis and did not undergo precocious oocyte maturation, suggesting that EGA occurs independently of cell cycle events.

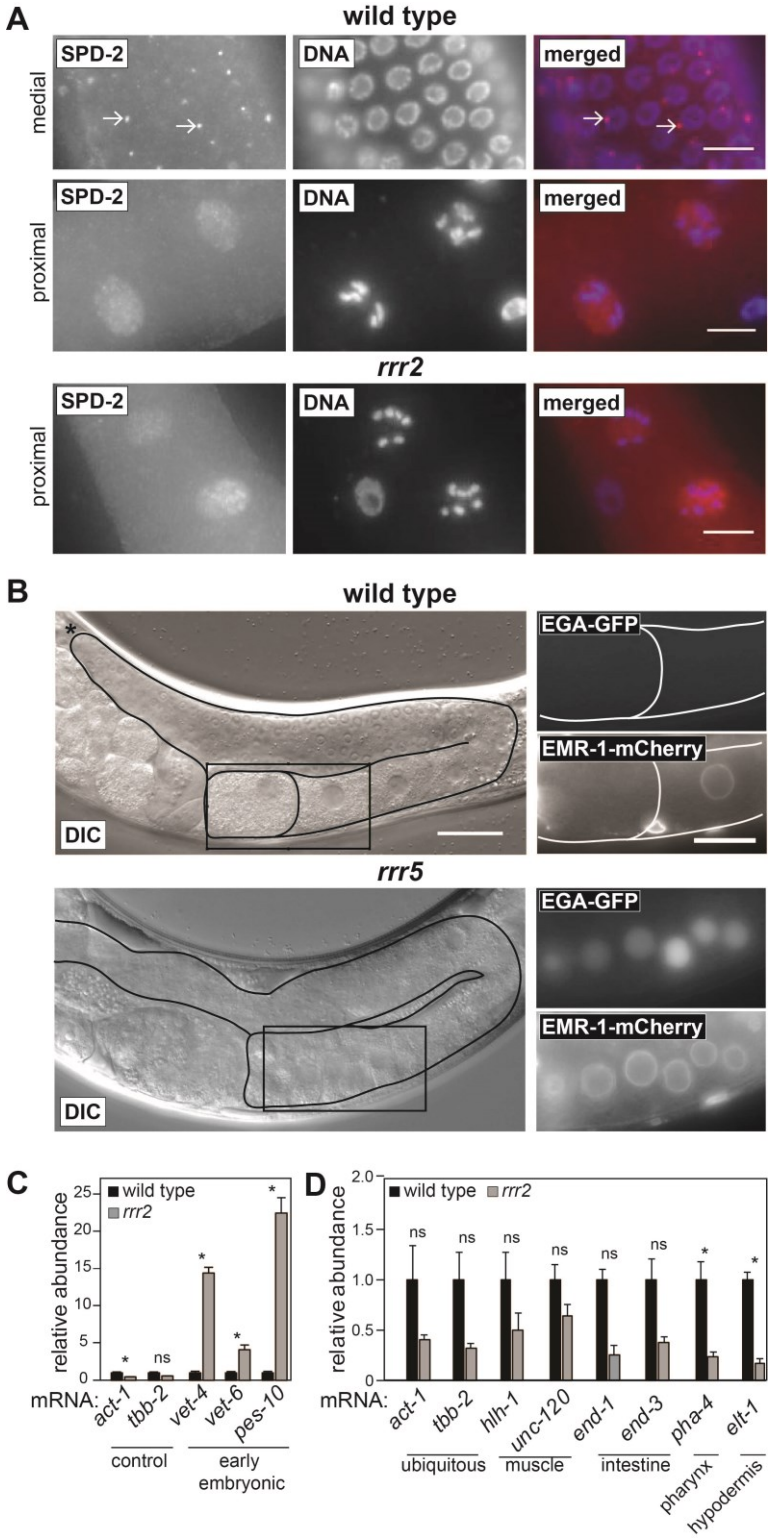

A
wild type
SPD-2
DNA
merged
medial
proximal
rrr2
B
EGA-GFP
EMR-1-mCherry
DIC
rrr5
C
wild type
rrr2
relative abundance
mRNA:
act-1
tbb-2
vet-4
vet-6
pes-10
control
early embryonic
D
hlh-1
unc-120
end-1
end-3
pha-4
elt-1
ubiquitous
muscle
intestine
pharynx
hypodermis
ns

Figure 3.2: Precocious EGA is cell cycle-independent and limited to early expressed embryonic genes in the newly identified mutants

A. Fluorescent micrographs of wild type and *rrr2* mutant gonads, immunostained for the centrosomal component SPD-2 and DAPI. The medial view shows the presence of centrosomes in distal meiotic germ cells in wild type. The proximal views show wild type and *rrr2* mutant oocytes, where no SPD-2 staining could be detected, indicating the elimination of centrosomes. Scale bars: 10 µm. **B.** DIC micrographs of live animals, either wild type or *rrr5*. Scale bar: 30 µm. Boxed areas are enlarged on the right and show EGA-GFP reporter expression or expression of an EMR-1-mCherry transgene to visualize nuclear envelopes. In wild type, oocytes show EMR-1 localization to the nuclear envelope, except for the -1 oocyte that underwent nuclear envelope breakdown. *rrr5* mutant oocytes, which express the EGA-GFP reporter, have intact nuclear envelopes. Scale bar: 20 µm. **C.** Detection of EGA transcripts by qRT-PCR. "Early embryonic": mRNAs, which are expressed immediately after the onset of EGA. Each bar represents the mean of three independent biological replicates. Error bars represent the standard error of the mean (SEM). The significance of the differences has been calculated with the Student's t-test: "*", $p<0.05$; "*n.s.*", not significant. **D.** Detection of somatic lineage-specific transcripts by qRT-PCR. The respective tissues and representative transcripts are indicated. Each bar represents the mean of three independent biological replicates. Error bars and p-values calculated as in C.

Mutants express early embryonic transcripts, but not lineage-specific transcripts

To characterize the extent of embryonic transcription in mutant gonads, we examined the levels of the endogenous *vet-4* and additional embryonic transcripts using reverse transcription and quantitative PCR (RT-qPCR). We found that early expressed embryonic transcripts (Baugh et al. 2003), including *vet-4*, were highly upregulated in *rrr2* mutant gonads compared to wild type (Figure 3.2C). Somatic lineage-specific transcripts which are expressed later during embryonic development, however, are not upregulated in *rrr2* mutant gonads (Figure 2D). These findings indicate that *rrr2* mutant gonads specifically express early embryonic transcripts, but not transcripts involved in somatic differentiation. This observation is in contrast to what has been shown for *gld-1* and *lin-41* mutant gonads, where the maternal transcription factor PAL-1, which drives muscle differentiation in embryos (Hunter and Kenyon 1996), leads to differentiation of the reprogrammed germ cells into muscles (Tocchini et al. 2014; Ciosk et al. 2006). Thus, the mutant germ cells appear to undergo only a first reprogramming step, resulting in the expression of early embryonic genes, but do not undergo a full reprogramming to pluripotency. Being able to separate EGA from an additional reprogramming step, leading to teratoma formation in *gld-1* and *lin-41* mutants, provides the unique opportunity to study EGA repression separately, uncoupled from additional phenotypes.

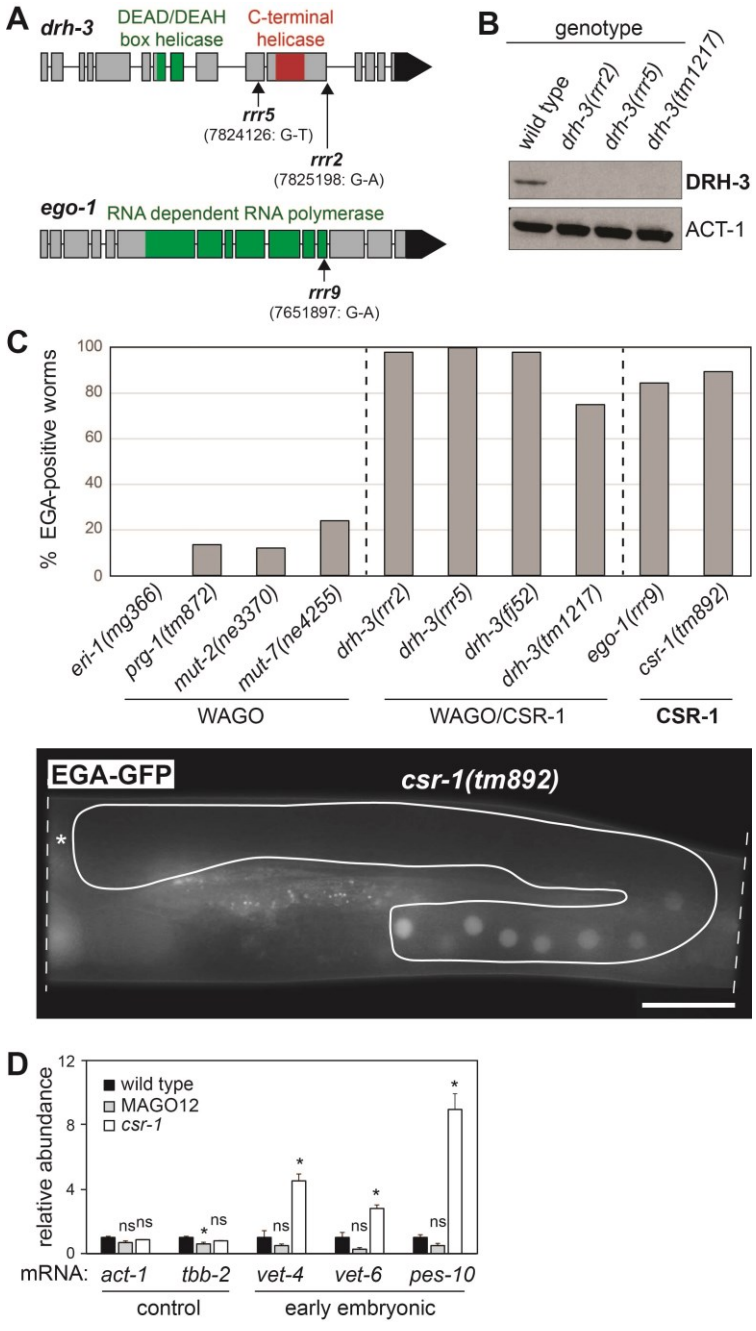

A
drh-3
DEAD/DEAH box helicase
C-terminal helicase
rrr5
(7824126: G-T)
rrr2
(7825198: G-A)
ego-1
RNA dependent RNA polymerase
rrr9
(7651897: G-A)
B
genotype
wild type
drh-3(rrr2)
drh-3(rrr5)
drh-3(tm1217)
DRH-3
ACT-1
C
% EGA-positive worms
100
80
60
40
20
0
eri-1(mg366)
prg-1(tm872)
mut-2(ne3370)
mut-7(ne4255)
drh-3(rrr2)
drh-3(rrr5)
drh-3(fj52)
drh-3(tm1217)
ego-1(rrr9)
csr-1(tm892)
WAGO
WAGO/CSR-1
CSR-1
EGA-GFP
csr-1(tm892)
*
D
relative abundance
12
8
4
0
wild type
MAGO12
csr-1
ns
*
mRNA:
act-1
tbb-2
vet-4
vet-6
pes-10
control
early embryonic

Figure 3.3: EGA repression in the germline is mediated *via* the CSR-1 pathway

A. Schematic view of the *drh-3* and *ego-1* gene structures with highlighted known protein domains and their putative functions. Arrows indicate the locations of the identified mutations. The *drh-3(rrr2)* mutation is intronic, likely leading to a splice-donor variant. The *drh-3(rrr5)* mutation is located in exon 10 is causing a precocious stop codon (Q647>Stop). The *ego-1(rrr9)* mutation is located in exon 12 and leads to a nonsynonymous amino acid change (A1271>V). **B.** Western blot analysis of wild type and different *drh-3* mutant protein lysates probed for DRH-3 and ACT-1 (as a loading control). No full-length DRH-3 protein could be detected in the *drh-3(rrr2)* and *drh-3(rrr5)* protein lysates. *Drh-3(tm1217)* is a putative null allele (containing a 482 bp deletion) which was used as a negative control. *Drh-3* homozygous populations were obtained by worm sorting. **C.** Detection of EGA-GFP reporter expression in mutants belonging to the WAGO or CSR-1 pathway, as well as shared components of both pathways. Each bar represents the percentage of animals that express the reporter in germ cells. N>30. Fluorescent micrograph of a live *csr-1(tm892)* mutant, grown at 25°C, expressing the EGA-GFP reporter abnormally in developing oocytes. Scale bar: 40µm. **D.** Detection of EGA transcripts in wild type, a strain containing mutations in all 12 WAGOs (MAGO12) and *csr-1(tm892)* mutants by qRT-PCR. Each bar represents the mean of three independent biological replicates. Error bars represent the SEM. The significance of the differences has been calculated with the Student's t-test: "*", $p<0.05$; "*n.s.*", not significant.

The CSR-1 pathway functions in EGA repression

We mapped the mutants to *drh-3* (*dicer-related-helicase-3,* alleles *rrr2/rrr5*) and to *ego-1* (*enhancer-of-GLP-ONE-1,* allele *rrr9*) (Figure 3.3A and Figure S3.1), both functioning in small non-coding RNA pathways. All mutants seem to behave like molecular nulls, as they display fully penetrant sterility, which was reported earlier for putative *drh-3* and *ego-1* null alleles (Duchaine et al. 2006; Smardon et al. 2000). In addition, examining DRH-3 protein levels in *drh-3(rrr2/rrr5)* mutants by western blot, did not reveal the presence of the protein in neither of the two mutants (Figure 3.3B). In the *C. elegans* germline, two main endo-siRNA pathways have been reported to function, each employing different Argonautes: the WAGO (*worm argonaute*) pathway and the CSR-1 (*chromosome-segregation and RNAi-deficient-1*) pathway (Gu et al. 2009; Claycomb et al. 2009). These Argonautes bind specific classes of small RNAs, which are 22 nucleotides in length with a guanosine at the 5'end (22G RNAs)(Gu et al. 2009). 22G RNAs are produced by RNA-dependent RNA polymerases (RdRPs) and act as secondary effectors of the endo-siRNA pathways. DRH-3 is a core component present in all RdRP complexes and therefore important for the biogenesis of all classes of 22G-RNAs (Gu et al. 2009). EGO-1 functions redundantly with another RNA polymerase, called RRF-1, in the production of WAGO 22G-RNAs (Gu et al. 2009). For the production of CSR-1 22G RNAs, EGO-1 alone is required (Claycomb et al. 2009), suggesting that loss of the CSR-1 pathway function is responsible for the EGA phenotype. To confirm this hypothesis, we decided to examine the expression of the EGA reporter in mutants that are specific for either WAGO or CSR-1 pathway functions. Different

primary siRNA-pathways function upstream of the WAGO pathway and use the RdRP complex for the production of secondary 22G RNAs to amplify their response. They include the maternal ERGO-1 and spermatogenesis-specific ALG-3/4 26G RNA pathways (Vasale et al. 2010; Han et al. 2009; Conine et al. 2010), as well as the piRNA pathway (Bagijn et al. 2012; Lee et al. 2012; Luteijn et al. 2012). To address if the EGA reporter expression is the result of disrupting the 26G RNA pathways, we looked at EGA expression in *eri-1(mg366)* mutant germlines lacking 26G RNAs (Han et al. 2009). To test the involvement of the piRNA pathway in EGA repression, we used *prg-1(tm892)* mutants, lacking most piRNAs (Wang and Reinke 2008). Neither *eri-1* nor *prg-1* mutants showed EGA reporter expression in the germline (Figure 3.3C). In addition, loss of *mut-2* and *mut-7,* essential components of the mutator foci which are required for WAGO 22G RNA amplification (Phillips et al. 2012; Gu et al. 2009; Zhang et al. 2011) did not cause a robust germline-EGA phenotype. In contrast, a high percentage of *csr-1(tm892)* mutants showed EGA-GFP expression in the germline, comparable to the identified *ego-1(rrr9)* mutant (Figure 3.3C). In addition, we tested putative *drh-3* null alleles for EGA-GFP expression. Those mutants showed a similar percentage of germline-EGA expressing worms to the *drh-3(rrr2)* and *drh-3(rrr5)* mutants, which further suggested that our *drh-3* mutants behave like null mutants. To verify the results of our EGA reporter assay also for endogenous embryonic genes we used qRT-PCR. MAGO12 is a strain that combines mutants of all 12 WAGO genes (Gu et al. 2009) and can therefore be seen as a complete loss-of function of the WAGO pathway activity. We determined early embryonic transcript levels in this mutant and in the *csr-1(tm892)* mutant, where the loss of the only Argonaute in this pathway abolishes the function. We found that the embryonic transcripts we tested were highly upregulated in *csr-1* compared to wild type gonads, whereas no changes appeared in MAGO12 mutant gonads (Figure 3.3D). Together, these results suggest that it is the disruption of the CSR-1 endo-siRNA pathway that determines the germline-EGA phenotype.

A transcriptional regulation of the EGA reporter

Post-transcriptional gene regulation via the 3' UTR of an mRNA is a major mode of gene regulation in the germline (Merritt et al. 2008). We therefore wanted to examine if the EGA reporter is regulated *via* its 3'UTR. For the reporter we used the endogenous promoter and 3'UTR sequences of the embryonic gene *vet-4*. To distinguish between a transcriptional, promoter-dependent, versus a post-transcriptional regulation via the 3'UTR of our EGA reporter, we exchanged the gene-specific *vet-4* 3'UTR to an unregulated *tbb-2* 3'UTR. Exchanging the 3'UTR still resulted in the expression of the reporter in mutant germ cells, but not in the wild type ones (Figure 3.4A). We concluded that the regulation of our reporter

is dependent on its promoter. This suggested a transcriptional regulation of the reporter and likely also of the endogenous embryonic genes.

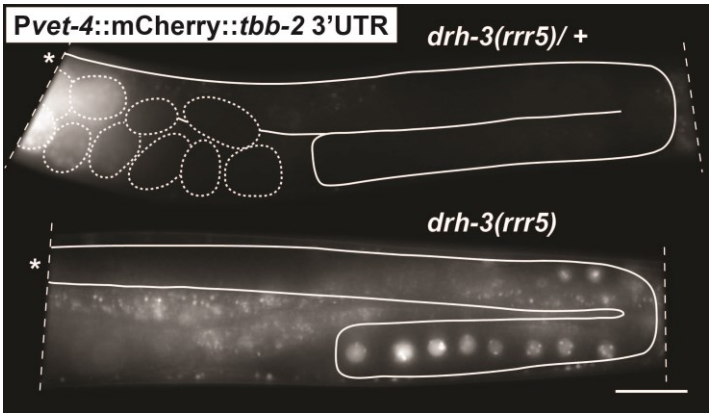


Figure 3.4: Expression regulation of the EGA reporter is promoter-dependent

Fluorescent micrographs of live animals expressing the *Pvet-4::mCherry::tbb-2 3'UTR* reporter. *Drh-3(rrr5)* heterozygotes express the reporter specifically in embryos, whereas *drh-3(rrr5)* homozygotes express the reporter abnormally in developing oocytes. Scale bar: 40 µm.

Histone protein depletion does not cause EGA

The above finding of the transcriptional regulation of the EGA reporter suggested that the loss one of the reported functions of the CSR-1 pathway in transcription or chromatin regulation might cause its precocious expression in germ cells. Among the functions potentially affecting chromatin formation, is the suggested role of CSR-1 in histone mRNA maturation (Avgousti et al. 2012). This study reported a function for the CSR-1 pathway in the processing of histone pre-mRNAs to mature mRNAs via binding of CSR-1 in an *ego-1*-dependent manner to histone pre-mRNA 3'UTRs and inducing their cleavage. Moreover, they report a severe depletion of histone proteins in CSR-1 pathway mutants resulting from the defects in histone pre-mRNA processing, which potentially results in a decreased histone incorporation into chromatin. We examined if our *drh-3* mutants show such a decrease in overall histone protein levels by western blot analysis. However, a decrease in histone H2B levels in our *drh-3* mutants and in the putative *drh-3(tm1217)* null mutant compared to wild type animals could not be detected (Figure 3.5A). It was unexpected that we did not observe a decrease in histone H2B levels in *drh-3(tm1217)* mutants, as this was reported before (Avgousti et al. 2012). We therefore wanted to interfere with histone protein production using a second approach. CDL-1 is the *C. elegans* histone stem-loop binding protein (SLBP) and it is required for the production of histone proteins (Keall et al. 2007). Consequently, RNAi knockdown of

CDL-1 leads to a severe depletion of histone proteins (Keall et al. 2007). If the reduction of histone protein levels is causing the germline-EGA, then the depletion of CDL-1 should give the same phenotype. We therefore examined the expression of our EGA reporter in *cdl-1*-depeted animals. EGA reporter expression in germ cells of *cdl-1*-depeted animals could, however, not be observed (Figure 3.5B). In sum, we did not detect a depletion of histone proteins in our *drh-3* mutants and also interfering with histone protein production did not cause the EGA phenotype. We conclude that the depletion of histone proteins is unlikely to be the cause for EGA misregulation.

The CSR-1 catalytic slicer activity is required for EGA repression

Recently, CSR-1 has been reported to degrade maternal mRNAs encoding proteins with essential functions in the early embryo via its slicing activity (Gerson-Gurwitz et al. 2016). We therefore wanted to test if the slicing function contributes to embryonic gene repression. We performed qRT-PCR analysis on gonads for embryonic transcipts using a CSR-1 slicer-dead mutant, containing a single-copy integrated FLAG::CSR-1 construct with the DDH catalytic slicer residues mutated to AAA in the *csr-1(tm892)* mutant background (Figure 3.5C). The CSR-1 slicer-dead mutant revealed upregulation of the tested embryonic transcripts to similar levels as detected in *csr-1(tm892)* null mutants (Figure 3.5C). This indicated that the slicer activity is required for EGA repression. The importance of such a post-transcriptional regulation by CSR-1 was surprising, because our previous results suggested a transcriptional regulation of embryonic genes, indicated by the transcriptional regulation of the EGA reporter. How can we explain the importance of a post-transcriptional function for the transcriptional repression of embryonic genes? One explanation is that CSR-1 does not target embryonic transcripts directly, but controls the transcription of embryonic genes by regulating upstream factors *via* the slicing activity. This would mean that embryonic genes are not among the CSR-1 target genes, which is expected given that most CSR-1 targets are germline-expressed genes. To test this, we used a list of CSR-1 targets, which is based on the presence of gene-specific complementary 22G RNAs that associate with CSR-1 (Claycomb et al. 2009). Using transcriptome profiling data from staged *C. elegans* embryos (Baugh et al. 2003), we identified a group of 446 early embryonic genes. Embryonic transcription in *C. elegans* starts at the 4-cell stage (Seydoux and Fire 1994). Therefore, we defined the early embryonic genes as genes, which are upregulated at least 2-fold from the 4-cell to the 8-cell stage and from the 8-cell to the 15-cell stage, but are not present at the 4-cell stage, and are therefore not maternally deposited. Our analysis showed that the majority of early embryonic genes are not CSR-1 target genes (Figure 3.5D). Among the embryonic genes, which we tested previously by qRT-PCR only one, *vet-6*, appears to be a CSR-1 target. In sum, it seems unlikely that embryonic genes

are directly targeted by CSR-1 for degradation. We suggest that CSR-1 regulates *via* its slicing activity the abundance of a maternal mRNA, encoding for example a transcription factor or chromatin regulator, which contributes to the transcriptional regulation of embryonic genes.

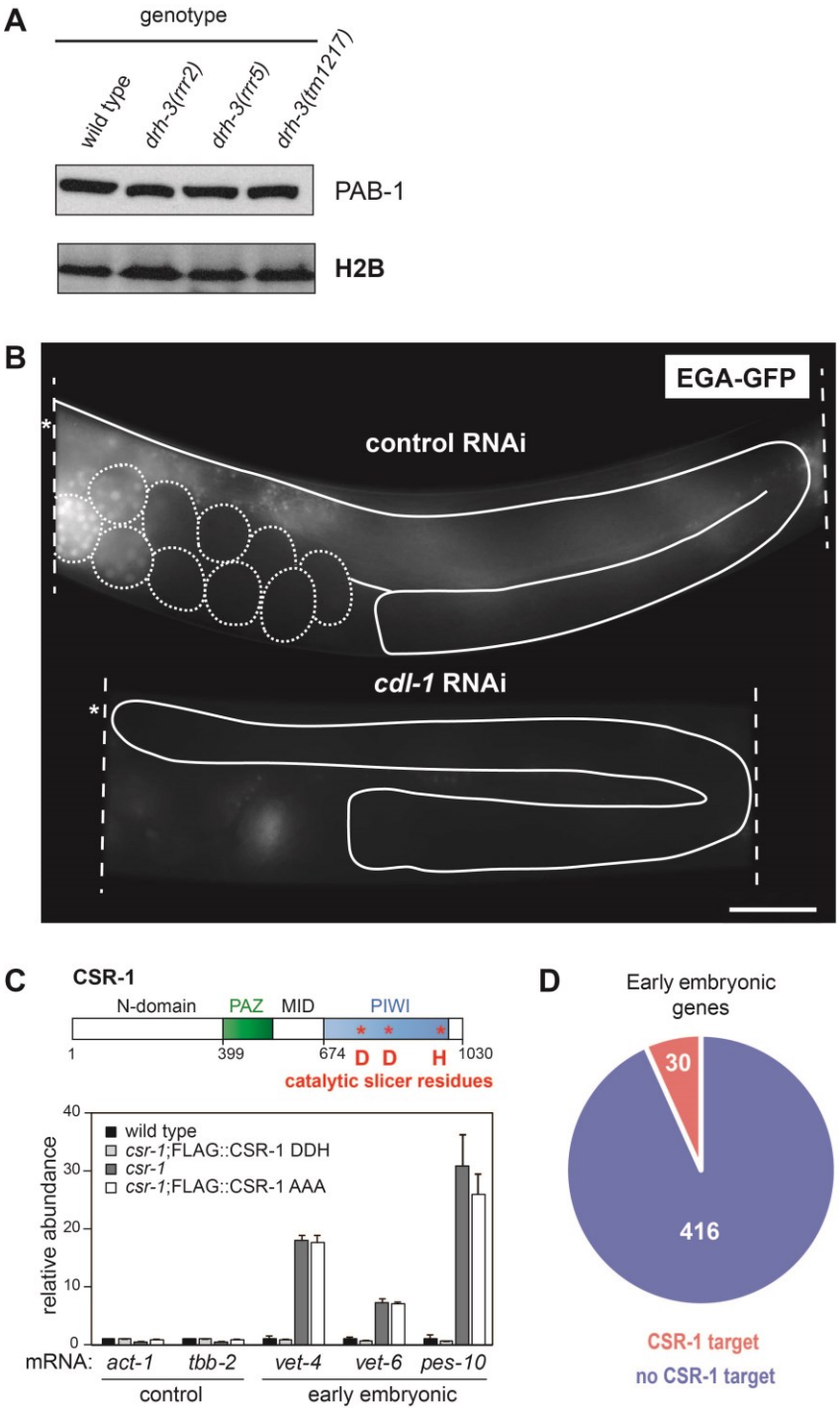


Figure 3.5: The depletion of histones does not cause a precocious EGA, whereas inactivation of the CSR-1 slicer function does

A. Western blot analysis of wild type and different *drh-3* mutant protein lysates probed for H2B and PolyA-binding protein-1 (PAB-1, as a loading control). A decrease in H2B protein levels in *drh-3(tm1217)* mutants, as described before in (Avgousti et al. 2012), or in our newly identified *drh-3(rrr2/rrr5)* mutants, could not be detected. **B.** Fluorescent micrographs of live animals expressing the EGA-GFP transgene after control or *cdl-1* RNAi treatment. *Cdl-1* RNAi treated animals were sterile, producing only unfertilized oocytes. Like control RNAi treated animals, also *cdl-1* RNAi treated animals did not show EGA-GFP expression in germ cells (1 out of 45 sterile animals). Scale bar: 30 µm. **C.** Schematic view of the CSR-1 protein with highlighted known protein domains. In red indicated are the DDH residues that confer the catalytic slicer function. In CSR-1 slicer-dead mutants DDH residues are mutated to AAA. Detection of EGA transcripts by qRT-PCR. *csr-1(tm892)* null mutants and *csr-1*; FLAG::CSR-1 AAA (slicer-dead) mutants show upregulation of early embryonic genes. Each bar represents the mean of three independent biological replicates. Error bars represent the SEM. **D.** Among a group of 446 early expressed embryonic genes, only 30 are considered to be CSR-1 targets, based on the presence of antisense 22G RNAs associated with CSR-1 in pull-down experiments.

3.1.2 Discussion

A model to study EGA repression

We report here the identification of mutants undergoing a precocious EGA in germ cells without a re-entry into the cell cycle. Earlier studies of mutants showing EGA in the germline suggested that this is a cell-cycle dependent event, allowing the transcription of embryonic genes only after a re-entry into mitosis (Biedermann et al. 2009; Tocchini et al. 2014) or after a precocious oocyte maturation (Allen et al. 2014). Being able to separate EGA from cell cycle events provides the unique opportunity to study the transcriptional repression of embryonic genes in germ cells without complicating the analysis by additional phenotypes due to cell cycle misregulation.

In *C. elegans* embryos, repression of the basal transcription machinery has been implicated in EGA control. In one- and two-cell stage embryos, TAF-4, a TFIID subunit, is bound by the proteins OMA-1/2 and thereby sequestered in the cytoplasm to repress RNA Pol II-dependent transcription (Guven-Ozkan et al. 2008). In the four-cell stage embryo, when EGA onset occurs, the degradation of OMA-1/2 leads to the release of TAF-4 into the nucleus, where TFIID assembly can occur and transcription start. How transcriptional repression is mediated in germ cells remains unclear. Germ cells are transcriptionally active until diakinesis, when transcription is shut down by an unknown mechanism. The expression of embryonic genes in the oocytes of CSR-1 pathway mutants suggests that these are transcriptionally active. This is

supported by the finding that the EGA reporter is transcriptionally regulated. However, a transcriptional activity of the mutant oocytes has to be confirmed further. Transcriptional activity can be monitored by the phosphorylations occurring at the C-terminal domain (CTD) on RNA Polymerase II (RNA-Pol II). During transcription initiation the CTD is phosphorylated on serine 5 (PSer5); during elongation the phosphorylation shifts to serine 2 (PSer2). Thus, the transcriptional activity can be monitored using antibodies specific to PSer5 and PSer2 (Seydoux and Dunn 1997). By performing immunostainings with these antibodies, we can therefore address the question if the mutant oocytes show globally active transcription. In addition, the transcription should be dependent on other factors involved in Pol-II mediated transcription, such as components of the pre-initiation complex (PIC) or the kinases phosphorylating the serines on the Pol II CTD. The dependence of the expression of the EGA reporter on these factors could be answered by simultaneous knockdown with CSR-1 pathway components. If the mutant oocytes show indeed active transcription, then further questions include if they remain transcriptionally active, despite the entry into diakinesis, or if they undergo *de novo* Pol II activation.

Reported CSR-1 functions and their contribution to EGA repression

The CSR-1 pathway has been implicated in many different functions in the germline. In our study, we addressed the possible contribution of some of them to EGA control. First, we tested if the suggested function of CSR-1 in histone mRNA maturation (Avgousti et al. 2012) could be required to repress a precocious EGA in oocytes. However, we could not detect a global decrease in histone protein levels in our *drh-3* mutants. Moreover, the depletion of CDL-1, the *C. elegans* stem-loop binding protein, which leads to a drastic decrease in histone proteins (Keall et al. 2007), did not cause a precocious EGA. Therefore, histone protein depletion does likely not account for the EGA phenotype. A reported negative regulatory role of CSR-1 in the FBF-1 pathway (Friend et al. 2012) is equally unlikely to contribute to EGA repression. FBF-1 is only expressed in mitotic germ cells (Crittenden et al. 2002) and thus not present in developing oocytes, where the precocious EGA occurs. In addition, this function appears to be small RNA-independent, as the defects in translational repression in *csr-1* mutants, were not (or much less) observed in mutants of the RdRP complex factors DRH-3 and EGO-1. Though, the role of CSR-1 in EGA repression is likely small RNA-dependent, as precocious EGA is also observed in *drh-3* and *ego-1* mutants. Another suggested activity of CSR-1 is the contribution to chromatin organization by regulating the incorporation of the centromere-specific histone H3 protein (CENP-A) and by regulating chromatin modifiers or histone methyltransferases (Cecere et al. 2014; Claycomb et al. 2009). However, the contribution of CSR-1 to this function remains correlative and an interaction of CSR-1 with chromatin remodelers has not been

published to date. Also, a suggested function of CSR-1 in P granule (the *C. elegans* germ granule) formation does likely not cause the precocious EGA. The loss of perinuclear localization of P granules has been reported for *csr-1* mutants and following the depletion of CSR-1 by RNAi knockdown (Claycomb et al. 2009; Campbell and Updike 2015). In contrast, P granules seem to be unaffected when the CSR-1 slicing activity is lost (Gerson-Gurwitz et al. 2016). As we observed a precocious EGA in CSR-1 slicer mutants, P granule loss does not appear to contribute to this phenotype, assuming that the CSR-1 slicer mutant we used in the experiments behaves in the same way as the published mutant. In general, it remains to be controversially discussed to which extent the loss of CSR-1 affects chromatin organization and P granule formation, as its loss is not accompanied by the strong phenotypes associated with loss of chromatin factors or P granules. In fact, *csr-1* mutant germ cells maintain germline characteristics and only subtle changes seem to occur in the transcriptome of *csr-1* mutants (Claycomb et al. 2009). Thus, the conditions leading to ectopic somatic gene expression in germ cells after global chromatin misregulation (Patel et al. 2012; Robert et al. 2014) or complete loss of P granules (Updike et al. 2014) differ considerably to the conditions where we see EGA. We observe EGA only in relatively normal looking oocytes, but not in strongly affected germ cells. Furthermore, CSR-1 is suggested to define transcribed genes in the germline as "self" and protect them from gene silencing by the piRNA pathway (Seth et al. 2013; Wedeles et al. 2013b). This protective function appears to be negligible for EGA repression, as we did not observe EGA in *prg-1* mutants, where piRNA function is abolished and silencing does not occur. Recently, the CSR-1 slicing activity has been found to be essential for early embryogenesis, potentially by regulating the abundance of maternally loaded transcripts into the embryo (Gerson-Gurwitz et al. 2016). It is important to note that the above mentioned CSR-1 functions do not take into account this new reported role of the CSR-1 slicing activity. Thus, it remains unknown if they are dependent on this endonucleolytic function. We found that the slicing activity is important to inhibit a precocious EGA in developing oocytes. This suggests that the slicing activity is already required to maintain functional oocytes, and is not only necessary for embryogenesis. How the slicing activity contributes to EGA repression on a mechanistic level is not clear. However, it seems rather unlikely that CSR-1 directly degrades embryonic transcripts, as these are presumably not expressed in germ cells and as the majority of embryonic genes are not CSR-1 targets. We suggest that CSR-1 regulates the abundance of a germline transcript that is important for the transcriptional repression of embryonic genes, for example by degrading transcripts encoding transcription factors or chromatin regulators. Another possibility is that CSR-1 is not specifically regulating the transcription of embryonic genes, but has an effect on transcription in general. One interesting observation indicating a global regulation of Pol II-mediated transcription, is the upregulation of the

positive Transcription Elongation Factor b (pTEFb) subunit, CIT-1.2, in CSR-1 slicer dead mutants (Gerson-Gurwitz et al. 2016). We attempt to identify a maternal target involved in transcription, by RNAi knockdown of candidate genes upregulated in the CSR-1 slicer-dead mutant and looking for a rescue of the EGA phenotype in *csr-1* mutants. However, their identification is complicated by the decreased knockdown efficiency in *csr-1* mutants, due to the function of CSR-1 also in the exogenous RNAi pathway. A similar RNAi screen for upstream candidates regulating EGA, based on up-and down-regulated genes in *drh-3* mutants from our RNA-sequencing data (Figure S3.2), was hampered by the same decreased exogenous RNAi activity.

Endogenous small RNA function in oocytes - a conserved mechanism?

Endo-siRNAs can be found in mammals, although they are not produced by RNA-dependent RNA polymerases (RdRPs) like they are in worms. RdRPs were likely present in the common eukaryotic ancestor, but have been lost in several eukaryotes, including vertebrates and insects (Zong et al. 2009). In mammals, endo-siRNAs are produced from long dsRNAs by Dicer processing (Suh and Blelloch 2011). Loss of Dicer in mouse oocytes results in meiotic arrest with spindle and chromosome segregation defects resulting in sterility (Murchison et al. 2007; Tang et al. 2007). Due to the function of Dicer in both, microRNA and endo-siRNA biogenesis, it was initially unclear which small RNA pathway function is important, or if both of them are. However, it has been found that microRNA function is globally suppressed in mouse oocytes (Suh et al. 2010; Ma et al. 2010), indicating that the Dicer mutant phenotype results from the loss of endo-siRNAs. The production of endo-siRNAs in mouse oocytes is predominantly performed by an oocyte-specific Dicer isoform, Dicer^{O}, which lacks the N-terminal DExD helicase domain (Flemr et al. 2013). The deletion phenotype of Dicer^{O} alone is identical to that of the deletion of both Dicer forms, the somatic and oocyte one. The deletion of AGO2, the only mammalian Argonaute with slicing activity, results in a very similar phenotype in mouse oocytes and it has been recently shown that the slicing activity is essential for endo-siRNA function in oocytes (Stein et al. 2015). In sum, a function for endo-siRNAs and their catalytic Argonautes in the female germline appears to be conserved. If the endo-siRNA pathway contributes in the mammalian germline, similarly to its function in the *C. elegans* germline, to EGA repression, remains to be addressed.

3.1.3 Supplemental figures

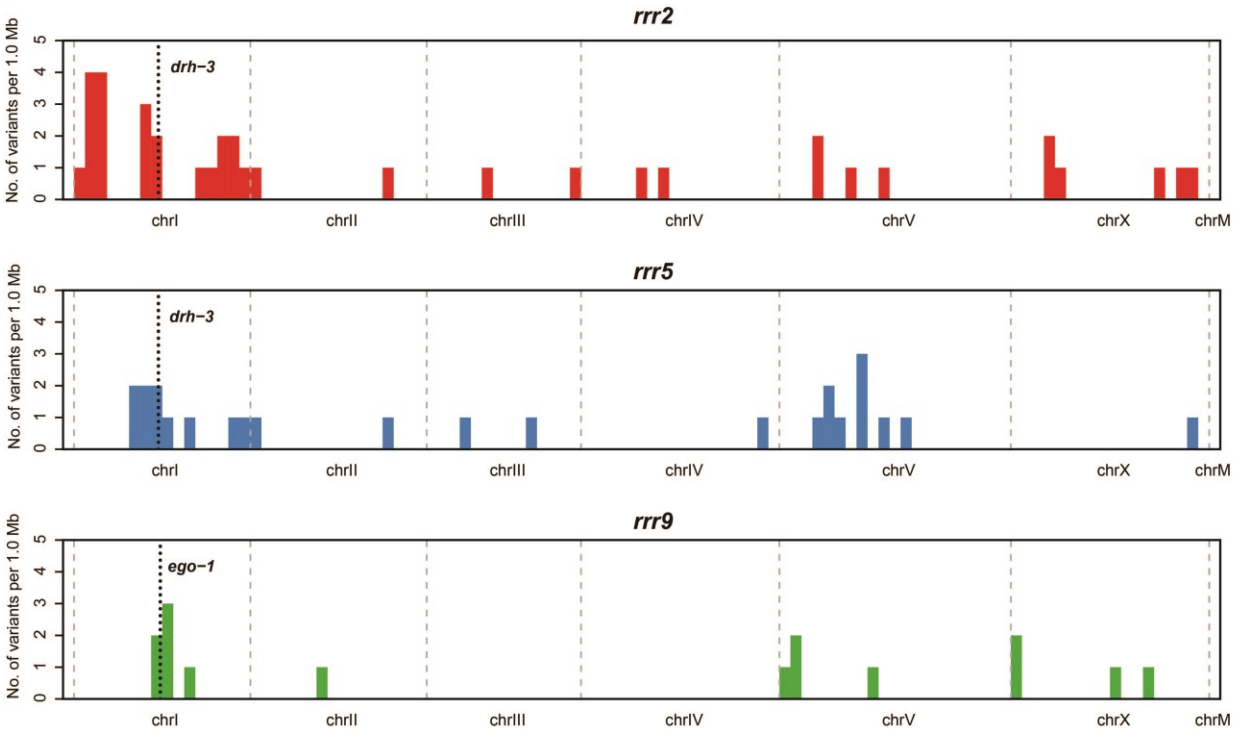


Figure S3.1: Single-nucleotide variants in the mutants induced by EMS-mutagenesis

After EMS-mutagenesis, all mutants were outcrossed against the wild type parental strain 5 to 8 times. Whole-genome-sequencing (WGS) revealed sequence variants clustering mostly on chromosome I. Numbers indicate chromosomes and “M” mitochondrial DNA. WGS data was filtered by "high quality SNVs (single nucleotide variants) of EMS-type not found in parent".

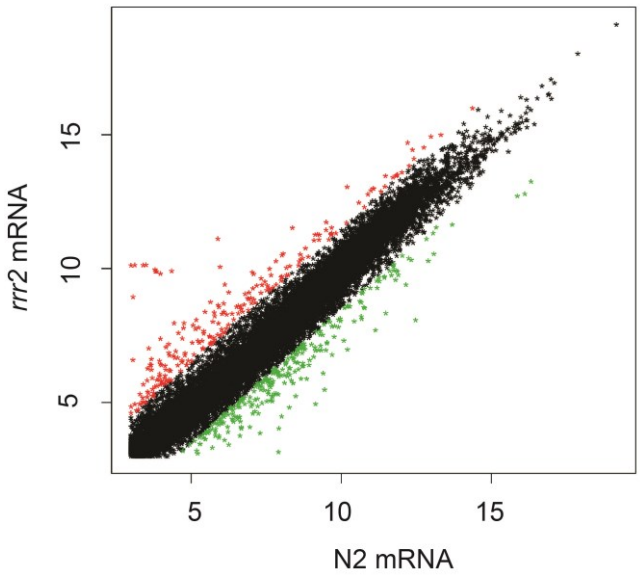


Figure S3.2: RNA-sequencing data *drh-3(rrr2)*

mRNA sequencing data from *drh-3(rrr2)* and wild type (N2) adult worms, performed in duplicates. The mean mRNA values were calculated and plotted along the y-axis *(drh-3(rrr2))* and x-axis (N2). In red marked are 220 mRNAs going up more than 1.5-fold (log scale) in *drh-3*. In green marked are 256 mRNAs going down more than 1.5-fold (log scale) in *drh-3.*

3.2 A common function for CSR-1 and LIN-41 in EGA repression?

with Pooja Kumari and Dimosthenis Gaidatzis

A previous study from our lab described the role of the RNA-binding protein LIN-41 in EGA repression in developing oocytes (Tocchini et al. 2014). To understand the mechanism underlying the LIN-41-mediated regulation of EGA, we performed LIN-41 immunoprecipitation (IP) experiments followed by mass-spectrometry analysis to identify interaction partners of LIN-41. Among others, CSR-1 was identified as an interaction partner, suggesting a common function of these proteins. Although LIN-41 appears to have an additional role in inhibiting the re-entry of germ cells into mitosis, CSR-1 and LIN-41 might function together to inhibit a precocious EGA. Here, we describe the interaction of the two proteins in more detail and present our approaches to understand the mechanism behind a potential cooperative regulation of EGA by LIN-41 and CSR-1.

3.2.1 Results

CSR-1 and LIN-41 Interaction

We discovered the *in vivo* interaction of LIN-41 with CSR-1 by performing LIN-41 IP-mass-spectrometry analysis (Figure 3.6A). Importantly, this interaction appeared to be RNA-independent as binding of the two proteins was consistent after RNase treatment. To confirm a direct interaction of the two proteins and to map the interacting domains, we used a cell-free protein expression system to perform Co-IP experiments. N-terminally tagged CSR-1 and LIN-41 proteins were expressed from plasmids under the control of an SP6 promoter in wheat-germ lysates (Figure 3.6B). These lysates contain all components necessary for transcription and translation, but are depleted of other mRNAs. Therefore, proteins expressed in the lysate can be directly used for Co-IP experiments without the need of further purification. The *in-vitro* produced proteins were allowed to form complexes and then tested for an interaction by pulling down one protein with the beads and looking for the presence of the interaction partner in the pull down after several washes. To detect the proteins, they were separated by SDS-PAGE and visualized with antibodies for the respective tags. Using this approach, we could confirm that CSR-1 binds to LIN-41 in an RNA-independent manner (Figure 3.6C).

LIN-41 belongs to the TRIM-NHL proteins, named after the presence of a TRIpartite Motif (TRIM) and a NCL-1/HT2A/LIN-41 (NHL) domain. The N-terminally located TRIM motif contains a RING domain, two B-

boxes and a coiled-coil domain (Figure 3.6D). In other TRIM-NHL proteins the TRIM motif has been suggested to function as an E3 ubiquitin ligase (Wulczyn et al. 2011). The TRIM motif is followed by a filamin domain with a so far unknown function in TRIM-NHL proteins (Wulczyn et al. 2011). Previous structure-function studies of LIN-41 showed that interfering with the RING or filamin domain function did not cause a precocious EGA in germ cells (Tocchini et al. 2014). The C-terminal part of LIN-41 contains the NHL domain, consisting of 6 NHL-repeats, which have been implicated in mRNA binding (Loedige et al. 2014). The introduction of a few point mutations in the RNA binding surface of the NHL domain led to precocious EGA in the germline (Tocchini 2015). Therefore, a functional NHL domain seems to be required for the LIN-41-mediated EGA repression, suggesting the importance of its mRNA regulatory role. We tested if the NHL domain is required for the binding to CSR-1. We could confirm the association of the LIN-41 NHL domain with CSR-1 in our Co-IPs (Figure 3.6E). As a positive control for our binding assay served the LIN-41 full-length protein, which was found to be associated with CSR-1 and as a negative control served GFP, which was not enriched in the CSR-1 IPs.

Next, we set out to identify the CSR-1 domain required for the interaction with LIN-41. CSR-1, similar to other Argonaute proteins, features the so-called N (N-terminal) domain, a PAZ (PIWI-Argonaute-Zwille) domain, a MID (middle) domain and a PIWI (P element-induced wimpy testis) domain (Figure 3.6D) (Swarts et al. 2014). The N domain has been shown in other Argonautes to play a role in target RNA cleavage and in unwinding the guide-passenger small RNA duplex. The PAZ domain forms a nucleic-acid binding structure and binds the 3'end of the guide RNA. The MID domain forms a nucleotide-binding pocket and interacts with the 5'end of the guide RNA and is possibly required to recognize specific 5'end bases via the "nucleotide specificity loop". Finally, the PIWI domain includes the catalytic slicer residues DDH, which are conserved in CSR-1 (Yigit et al. 2006). RNA cleavage mediated by this slicer residues is suggested to tune maternal transcript levels in the *C. elegans* germline to allow normal embryonic development (Gerson-Gurwitz et al. 2016). In addition, we found that the catalytic slicer activity of CSR-1 is important for EGA repression (Chapter 3.1). Apart from the small RNA binding and RNA-cleavage activities, the MID and PIWI domains have been also implicated in protein binding: CSR-1 has been suggested to bind the Pumilio-related RNA-binding protein FBF-1 via these domains (Friend et al. 2012). We therefore tested if the CSR-1 MID-PIWI domain can bind to the LIN-41-NHL domain. Such interaction was indeed found in our experiments using a MID-PIWI construct for the IP and detecting associated proteins (Figure 3.6F). In sum, the *in-vitro* binding studies, show an interaction of the CSR-1 MID-PIWI domain with the LIN-41 NHL domain.

A

Accession	Description	Enrichment in LIN-41 IP over control IP RNase – (fold change)	Enrichment in LIN-41 IP over control IP RNase + (fold change)
H2KZD5-CAEEL	CSR-1, isoform a	5	7.42

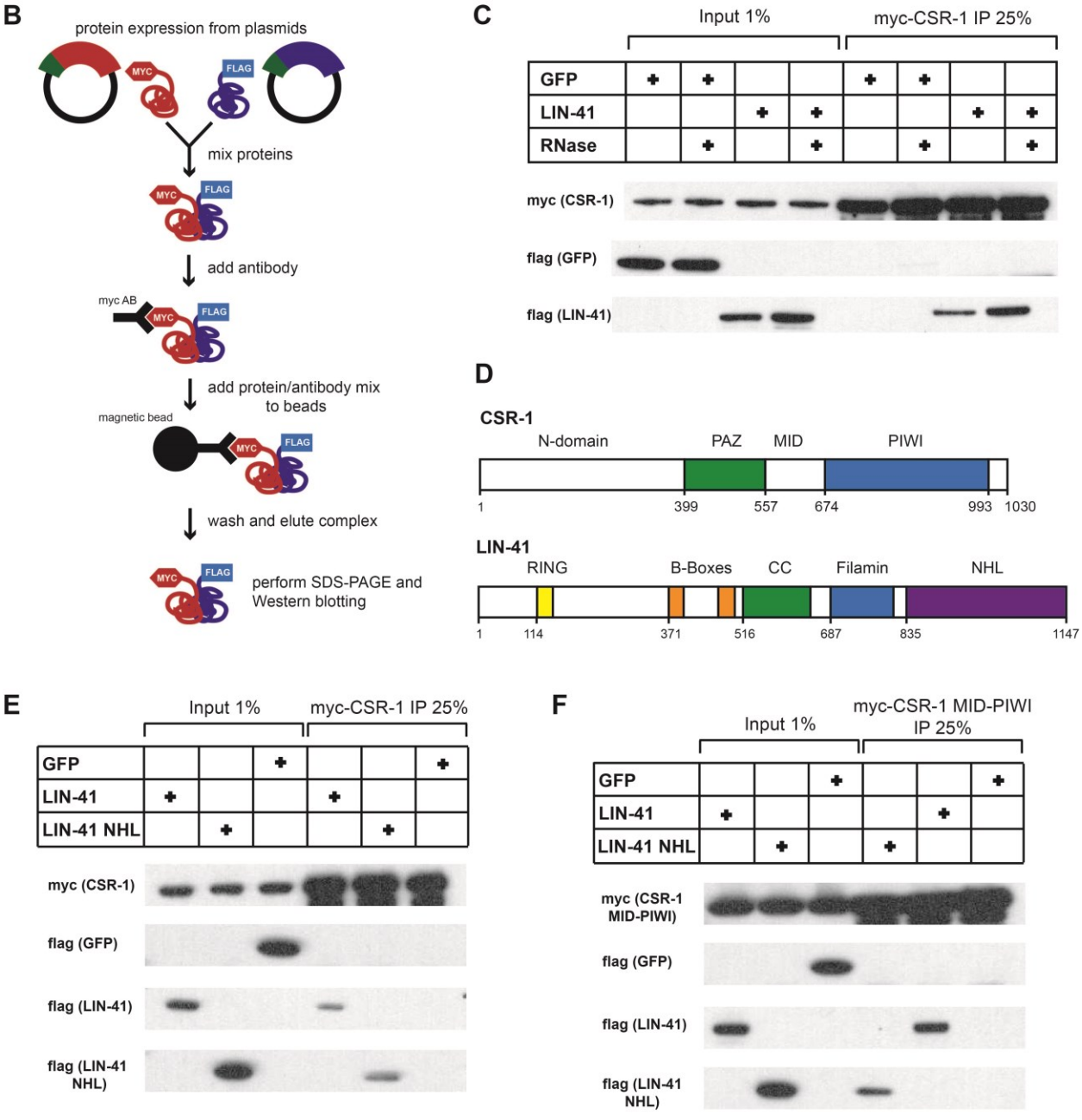


Figure 3.6: Interaction of CSR-1 and LIN-41

A. CSR-1 was identified in LIN-41 immunoprecipitation (IP) experiments followed by mass spectrometry analysis. CSR-1 was present in samples where RNA was present (RNase -) and where RNA was degraded (RNase +), indicating an RNA-independent interaction. **B.** Schematic of the co-immunoprecipitation experiments. Proteins were produced from plasmids in wheat-germ lysates and allowed to form complexes for 2 hours on ice. Anti-myc-tag antibodies were bound for 1 hour at 4°C before applying the protein/antibody mix to magnetic beads. Proteins were eluted from beads and loaded on a SDS gel. Western blot analysis was done using MYC and FLAG antibodies as indicated. **C.** Co-immunoprecipitations of myc-CSR-1 (full) and flag-LIN-41 (full). Flag-GFP was used as a negative control and experiments performed as described in B. "+" indicates the presence of the respective protein/RNase in the sample. LIN-41 co-immunoprecipitated with CSR-1 in an RNA-independent manner. **D.** CSR-1 and LIN-41 proteins with indicated known protein domains. **E.** Co-immunoprecipitations of myc-CSR-1 (full) and flag-LIN-41-NHL. Flag-GFP was used as a negative control, flag-LIN-41 (full) as a positive control. The LIN-41-NHL domain co-immunoprecipitated with CSR-1. **F.** Co-immunoprecipitations of myc-CSR-1-MID-PIWI and flag-LIN-41-NHL. Flag-GFP was used as a negative control, flag-LIN-41 (full) as a positive control. The LIN-41-NHL domain co-immunoprecipitated with the CSR-1-MID-PIWI domain.

Our primary goal is to make point mutations *in vivo* which will inhibit the binding of CSR-1 to LIN-41, but do not interfere with other functions of the proteins. Therefore, we aim to identify the amino acid residues that mediate the binding. To identify such amino acids, we made use of peptide microarray studies, performed by JPT Peptide Technologies. We provided the LIN-41 FIL-NHL domains, expressed and purified from insect cells. These proteins were labeled with a fluorophore and applied to a microarray consisting of 18 amino acid long overlapping peptides of CSR-1. Binding of the LIN-41 FIL-NHL domain to a CSR-1-peptide on the array was determined by the presence of a fluorescent signal, after washing excess un-bound protein off the microarray. The results were visualized in a heat-map diagram showing fluorescence intensities in a color-coded manner, from white (no binding) to red (strong binding) (Figure 3.7). As an internal process-control served normal human serum (NHS) and a murine anti-FLAG antibody, detected by secondary fluorescently-labelled anti-human and anti-mouse IgG antibodies. With this assay, the LIN-41 FIL-NHL binding to several CSR-1 peptides was detected. Peptide positions that showed a high fluorescent intensity and an intensity dependent on the concentration of the LIN-41 FIL-NHL protein applied to the array, were considered as peptides likely interacting with the LIN-41 FIL-NHL domain (Table 3.1). In addition, as the CSR-1 MID-PIWI domain was printed in duplicates on the array, the detection of identical peptides in both arrays gave more confidence on the specificity of the interaction. Using these filter criteria, we identified one peptide sequence (peptides 127 and 157) that we considered to be most likely mediating the interaction with the LIN-41 FIL-NHL domain. It has to be mentioned that the control serum and anti-FLAG antibody also led to increased fluorescence at the same peptide positions as the LIN-

41 FIL-NHL domain. Therefore, the identified peptide hits have to be considered with caution, because the peptides detected as interacting with the LIN-41 FIL-NHL domain might represent peptides that are more easily bound by other peptides in general. Despite these concerns, we looked at the most promising peptide hit in more detail. The peptides 126 and 157 share a 9 amino acid sequence with the peptides prior to them (peptides 125 and 156), which also showed an increased fluorescence in the array. This suggested that the overlapping 9 amino acid sequence may mediate the binding. To gain more information on this 9 amino acid sequence, we performed homology modeling of the *C. elegans* CSR-1 MID-PIWI domain based on the human Argonaute AGO2 structure. This modeling suggested that the 9 amino acid motif is part of an α-helix structure in the CSR-1 PIWI domain (data not shown). However, the motif is not conserved in between CSR-1 and AGO2 and hence it is difficult to predict the structural contribution of these residues. A deletion of the whole 9 amino acid motif would likely affect the CSR-1 structure and single residues to mutate are hard to predict. Due to the difficulties to predict the structural importance of the 9 amino acid motif and our concerns regarding the specificity of the interactions detected in the peptide microarray, we did not proceed to mutate this motif in *vivo* to look for a loss of the interaction of CSR-1 and LIN-41.


Human_IgG - Control-Spot
Human_IgG - Control-Spot
Human_IgG - Control-Spot
Mouse_IgG - Control-Spot
Mouse_IgG - Control-Spot
Mouse_IgG - Control-Spot
MNQKQNPRLALNIPGLEL - Peptide_001
ALNIPGLELSERTIPRHV - Peptide_002
SERTIPRHVVQMKLIDRQ - Peptide_003
VQMKLIDRQHNKEYILTT - Peptide_004
HNKEYILTTMSARGRGNR - Peptide_005
MSARGRGNRATKQKDNFI - Peptide_006
ATKQKDNFILLDILLKQW - Peptide_007
LLDILLKQWAAKKGQQNL - Peptide_008
AAKKGQQNLPAFAYDGAQ - Peptide_009
PAFAYDGAQSLFTLEGIS - Peptide_010
SLFTLEGISLMVDIKKED - Peptide_011
LMVDIKKEDALEIPELSN - Peptide_012
ALEIPELSNFLKDSISFL - Peptide_013
FLKDSISFLSGDLEISCE - Peptide_014
SGDLEISCEPDLEKPSFV - Peptide_015
PDLEKPSFVQTELNEWSD - Peptide_016
QTELNEWSDPRFYAYLDI - Peptide_017
PRFYAYLDIVTSQSAIRS - Peptide_018
VTSQSAIRSERYLSQSKG - Peptide_019
ERYLSQSKGLYVHTQSLE - Peptide_020
LYVHTQSLEELRVKWAVA - Peptide_021
ELRVKWAVAAKGIHKGCR - Peptide_022
AKGIHKGCRIVGANGPLP - Peptide_023
IVGANGPLPILELDPQST - Peptide_024
ILELDPQSTQVYASIPLS - Peptide_025
QVYASIPLSQMLQYAFPR - Peptide_026
QMLQYAFPRDFPPNRIVN - Peptide_027
DFPPNRIVNPNMKLQRAV - Peptide_028
PNMKLQRAVKLLLKDLKC - Peptide_029
KLLLKDLKCNPYYDDKQI - Peptide_030
NPYYDDKQIWATNTITVS - Peptide_031
WATNTITVSDVDYNAPKD - Peptide_032
DVDYNAPKDPEFRQKYPN - Peptide_033
PEFRQKYPNLKFPMLPAV - Peptide_034
LKFPMLPAVQCGTGPHKR - Peptide_035
QCGTGPHKRLMPLEYLKV - Peptide_036
LMPLEYLKVLPYQSIDRR - Peptide_037
LPYQSIDRRVLEEFELTP - Peptide_038
VLEEFELTPRANAPNERW - Peptide_039
RANAPNERWSTLQKHYDQ - Peptide_040
STLQKHYDQFGFNDQVMK - Peptide_041
FGFNDQVMKDFGVQICND - Peptide_042
DFGVQICNDPFNNVSEID - Peptide_043
PFNNVSEIDGERVLAPSV - Peptide_044
GERVLAPSVAYADPVHVD - Peptide_045
AYADPVHVDDEKRDWKAQ - Peptide_046
DEKRDWKAQDKKFVTPAT - Peptide_047
DKKFVTPATIDHLMFVLV - Peptide_048
IDHLMFVLVAGYTRTWDA - Peptide_049
AGYTRTWDADCDATKFVA - Peptide_050
DCDATKFVAKAFMQRCKD - Peptide_051
KAFMQRCKDKGMHIGSYS - Peptide_052
KGMHIGSYSMDQHNGERG - Peptide_053
MDQHNGERGSENFLTSVF - Peptide_054
SENFLTSVFKNLVTHPNY - Peptide_055
KNLVTHPNYRDSSFTPFV - Peptide_056
RDSSFTPFVLFISDDVPN - Peptide_057
LFISDDVPNIHECLKFEE - Peptide_058
IHECLKFEERMSDIPTQH - Peptide_059
RMSDIPTQHVLLKNVKKM - Peptide_060
VLLKNVKKMRDNIEKKSQ - Peptide_061
RDNIEKKSQGGRRAYDLT - Peptide_062
GGRRAYDLTLDNIVMKAN - Peptide_063
LDNIVMKANIKCGGLNYT - Peptide_064
IKCGGLNYTADIPRDLAC - Peptide_065
ADIPRDLACWNEVSTFVI - Peptide_066
WNEVSTFVIGMDVAHPDR - Peptide_067
GMDVAHPDRNAAREGNPS - Peptide_068
NAAREGNPSTVGLSCNSA - Peptide_069
TVGLSCNSAENPYSFIGD - Peptide_070
ENPYSFIGDFLYTDPRRE - Peptide_071
FLYTDPRREAIQDEILRK - Peptide_072
AIQDEILRKFTDQSVRNF - Peptide_073
FTDQSVRNFAEIRGFPKK - Peptide_074
AEIRGFPKKVIIFRDGVS - Peptide_075
VIIFRDGVSFGEETAALK - Peptide_076
FGEETAALKEVEIIEQTI - Peptide_077
EVEIIEQTIKTAAKSMGH - Peptide_078
KTAAKSMGHSDYAPKVLA - Peptide_079
SDYAPKVLAIVVKKRHHT - Peptide_080
IVVKKRHHTRFYAKGGHH - Peptide_081
RFYAKGGHHGNMPINPLP - Peptide_082
GNMPINPLPDTSVGGDIA - Peptide_083
DTSVGGDIAEYGKRQIFI - Peptide_084
EYGKRQIFIQAFRPVQGT - Peptide_085
QAFRPVQGTAKVPSFLVI - Peptide_086
AKVPSFLVIRDDEEVSDE - Peptide_087
RDDEEVSDEHVAKMVCAV - Peptide_088
HVAKMVCAVCSLHQLVNS - Peptide_089
CSLHQLVNSPTSIPTPVY - Peptide_090
PTSIPTPVYVAHELAKRG - Peptide_091
VAHELAKRGTGLYKAYRF - Peptide_092
TGLYKAYRFKNGELFDDW - Peptide_093
KNGELFDDWETLTTQLSY - Peptide_094
ETLTTQLSYSTLDRLSKV - Peptide_095
LIN-41 FIL-NHL 1 µg/mL
LIN-41 FIL-NHL 0.2 µg/mL
control NHS+α-FLAG AB
TTQLSYSTLDRLSKVRVV - Peptide_096
EIDGERVLAPSVAYADPV - Peptide_097
PSVAYADPVHVDDEKRDW - Peptide_098
HVDDEKRDWKAQDKKFVT - Peptide_099
KAQDKKFVTPATIDHLMF - Peptide_100
PATIDHLMFVLVAGYTRT - Peptide_101
VLVAGYTRTWDADCDATK - Peptide_102
WDADCDATKFVAKAFMQR - Peptide_103
FVAKAFMQRCKDKGMHIG - Peptide_104
CKDKGMHIGSYSMDQHNG - Peptide_105
SYSMDQHNGERGSENFLT - Peptide_106
ERGSENFLTSVFKNLVTH - Peptide_107
SVFKNLVTHPNYRDSSFT - Peptide_108
PNYRDSSFTPFVLFISDD - Peptide_109
PFVLFISDDVPNIHECLK - Peptide_110
VPNIHECLKFEERMSDIP - Peptide_111
FEERMSDIPTQHVLLKNV - Peptide_112
TQHVLLKNVKKMRDNIEK - Peptide_113
KKMRDNIEKKSQGGRRAY - Peptide_114
KSQGGRRAYDLTLDNIVM - Peptide_115
DLTLDNIVMKANIKCGGL - Peptide_116
KANIKCGGLNYTADIPRD - Peptide_117
NYTADIPRDLACWNEVST - Peptide_118
LACWNEVSTFVIGMDVAH - Peptide_119
FVIGMDVAHPDRNAAREG - Peptide_120
PDRNAAREGNPSTVGLSC - Peptide_121
NPSTVGLSCNSAENPYSF - Peptide_122
NSAENPYSFIGDFLYTDP - Peptide_123
IGDFLYTDPRREAIQDEI - Peptide_124
RREAIQDEILRKFTDQSV - Peptide_125
LRKFTDQSVRNFAEIRGF - Peptide_126
RNFAEIRGFPKKVIIFRD - Peptide_127
PKKVIIFRDGVSFGEETA - Peptide_128
GVSFGEETAALKEVEIIE - Peptide_129
ALKEVEIIEQTIKTAAKS - Peptide_130
QTIKTAAKSMGHSDYAPK - Peptide_131
MGHSDYAPKVLAIVVKKR - Peptide_132
VLAIVVKKRHHTRFYAKG - Peptide_133
HHTRFYAKGGHHGNMPIN - Peptide_134
GHHGNMPINPLPDTSVGG - Peptide_135
PLPDTSVGGDIAEYGKRQ - Peptide_136
DIAEYGKRQIFIQAFRPV - Peptide_137
IFIQAFRPVQGTAKVPSF - Peptide_138
QGTAKVPSFLVIRDDEEV - Peptide_139
LVIRDDEEVSDEHVAKMV - Peptide_140
SDEHVAKMVCAVCSLHQL - Peptide_141
CAVCSLHQLVNSPTSIPT - Peptide_142
VNSPTSIPTPVYVAHELA - Peptide_143
SIPTPVYVAHELAKRGTG - Peptide_144
EIDGERVLAPSVAYADPV - Peptide_145
PSVAYADPVHVDDEKRDW - Peptide_146
HVDDEKRDWKAQDKKFVT - Peptide_147
KAQDKKFVTPATIDHLMF - Peptide_148
PATIDHLMFVLVAGYTRT - Peptide_149
VLVAGYTRTWDADCDATK - Peptide_150
WDADCDATKFVAKAFMQR - Peptide_151
FVAKAFMQRCKDKGMHIG - Peptide_152
CKDKGMHIGSYSMDQHNG - Peptide_153
SYSMDQHNGERGSENFLT - Peptide_154
ERGSENFLTSVFKNLVTH - Peptide_155
SVFKNLVTHPNYRDSSFT - Peptide_156
PNYRDSSFTPFVLFISDD - Peptide_157
PFVLFISDDVPNIHECLK - Peptide_158
VPNIHECLKFEERMSDIP - Peptide_159
FEERMSDIPTQHVLLKNV - Peptide_160
TQHVLLKNVKKMRDNIEK - Peptide_161
KKMRDNIEKKSQGGRRAY - Peptide_162
KSQGGRRAYDLTLDNIVM - Peptide_163
DLTLDNIVMKANIKCGGL - Peptide_164
KANIKCGGLNYTADIPRD - Peptide_165
NYTADIPRDLACWNEVST - Peptide_166
LACWNEVSTFVIGMDVAH - Peptide_167
FVIGMDVAHPDRNAAREG - Peptide_168
PDRNAAREGNPSTVGLSC - Peptide_169
NPSTVGLSCNSAENPYSF - Peptide_170
NSAENPYSFIGDFLYTDP - Peptide_171
IGDFLYTDPRREAIQDEI - Peptide_172
RREAIQDEILRKFTDQSV - Peptide_173
LRKFTDQSVRNFAEIRGF - Peptide_174
RNFAEIRGFPKKVIIFRD - Peptide_175
PKKVIIFRDGVSFGEETA - Peptide_176
GVSFGEETAALKEVEIIE - Peptide_177
ALKEVEIIEQTIKTAAKS - Peptide_178
QTIKTAAKSMGHSDYAPK - Peptide_179
MGHSDYAPKVLAIVVKKR - Peptide_180
VLAIVVKKRHHTRFYAKG - Peptide_181
HHTRFYAKGGHHGNMPIN - Peptide_182
GHHGNMPINPLPDTSVGG - Peptide_183
PLPDTSVGGDIAEYGKRQ - Peptide_184
DIAEYGKRQIFIQAFRPV - Peptide_185
IFIQAFRPVQGTAKVPSF - Peptide_186
QGTAKVPSFLVIRDDEEV - Peptide_187
LVIRDDEEVSDEHVAKMV - Peptide_188
SDEHVAKMVCAVCSLHQL - Peptide_189
CAVCSLHQLVNSPTSIPT - Peptide_190
VNSPTSIPTPVYVAHELA - Peptide_191
SIPTPVYVAHELAKRGTG - Peptide_192
GGSGGSGGGSDYKDDDDK - process-control
GGSGGSGGGSDYKDDDDK - process-control
LIN-41 FIL-NHL 1 µg/mL
LIN-41 FIL-NHL 0.2 µg/mL
control NHS+α-FLAG AB

Figure 3.7: Peptide microarray – heat map diagram

Heat map diagram of the incubation of the CSR-1 peptide microarray with LIN-41 FIL-NHL domain in 1 µg/mL or 0.2 µg/mL concentration or normal human serum (NHS) and mouse α-FLAG antibody as control. The mean fluorescence intensity values are shown as colors from white (no or low intensity), over yellow (middle intensity), to red (high intensity).

Peptide number	Sequence	Fluorescent intensity (dependent on LIN-41 concentration)	Detected in duplicates	CSR-1 location
2	ALNIFGLELSERTIFRHV	Medium (yes)	no	N-domain
18	PRFYAYLDIVTSQSAIRS	Medium (yes)	no	N-domain
26	QYYASIPLSQMLQYAFPR	Low (no)	no	N-domain
81	IVVKKRHHTRFYAKGGHH	Medium (yes)	no	PIWI domain
85	EYGKRQIFIQAFRPVQGT	Medium (yes)	no	PIWI domain
86	QAFRPVQGTAKVPSFLVI	Medium (yes)	no	PIWI domain
92	VAHELAKRGTGLYKAYRF	High (yes)	no	PIWI domain + C-terminal end
96	TTQLSYSTLDRLSKVRVV	Medium (yes)	no	C-terminal end
126	LRKFTDQSVRNFAEIRGF	Low (yes)	yes	PIWI domain
127	RNFAEIRGFPKKVIIFRD	High (yes)	yes	PIWI domain
133	VLAIVVKKRHHTRFYAKG	Medium (yes)	yes	PIWI domain
137	DIAEYGKRQIFIQAFRPV	Low (yes)	yes	PIWI domain
174	LRKFTDQSVRNFAEIRGF	Low (no)	yes	PIWI domain
175	RNFAEIRGFPKKVIIFRD	High (yes)	yes	PIWI domain
181	VLAIVVKKRHHTRFYAKG	Medium (yes)	yes	PIWI domain
185	DIAEYGKRQIFIQAFRPV	Low (yes)	yes	PIWI domain

Table 3.1: Peptide microarray – peptide hits

The table summarizes the peptide hits that were identified in the peptide microarray. The respective peptide numbers and sequences are indicated in the table. The peptide hits were analyzed for the intensity of the fluorescent signal, the dependence of the fluorescent signal on the concentration of LIN-41 FIL-NHL protein applied to the array and the presence of the peptide hit in the duplicates of the CSR-1 MID-PIWI domain printed on the array. In addition, information on the location of the peptide sequence in the CSR-1 protein is indicated. The most interesting peptide sequence (peptides 127 and 157) is indicated in red.

What could be the functional mechanism of the interaction?

As mentioned before, both, CSR-1 and LIN-41, inhibit a precocious expression of embryonic genes in the developing oocytes. This, together with the discovery of a direct physical interaction between the two proteins suggests a functional relationship in EGA repression. One possible relationship could be that CSR-1 controls LIN-41 expression, or *vice versa*. To address this possibility, we examined the expression of a GFP-tagged LIN-41 rescuing construct in *drh-3* mutants. LIN-41-GFP expression in *drh-3* mutants showed no obvious changes in expression levels or expression pattern compared to wild type (Figure 3.8A). Although we observed the GFP-LIN-41 expression in *drh-3* mutants, we also expect no obvious changes occurring in *csr-1* mutants, because CSR-1 pathway mutants display the same phenotypes. The expression of CSR-1 was observed with an endogenously-GFP tagged CSR-1. In *lin-41* heterozygote animals we observed a cytoplasmic, perinuclear localization of CSR-1, as reported before (Claycomb et al. 2009). In *lin-41* homozygous animals we observed a very similar CSR-1 expression and localization until the proximal germline where a teratoma is formed (Figure 3.8B). Although subtle expression changes might be difficult to detect in these mutant backgrounds, the overall expression levels and localization patterns of the proteins did not change. Thus, LIN-41 and CSR-1 do not appear to regulate each other's expression.

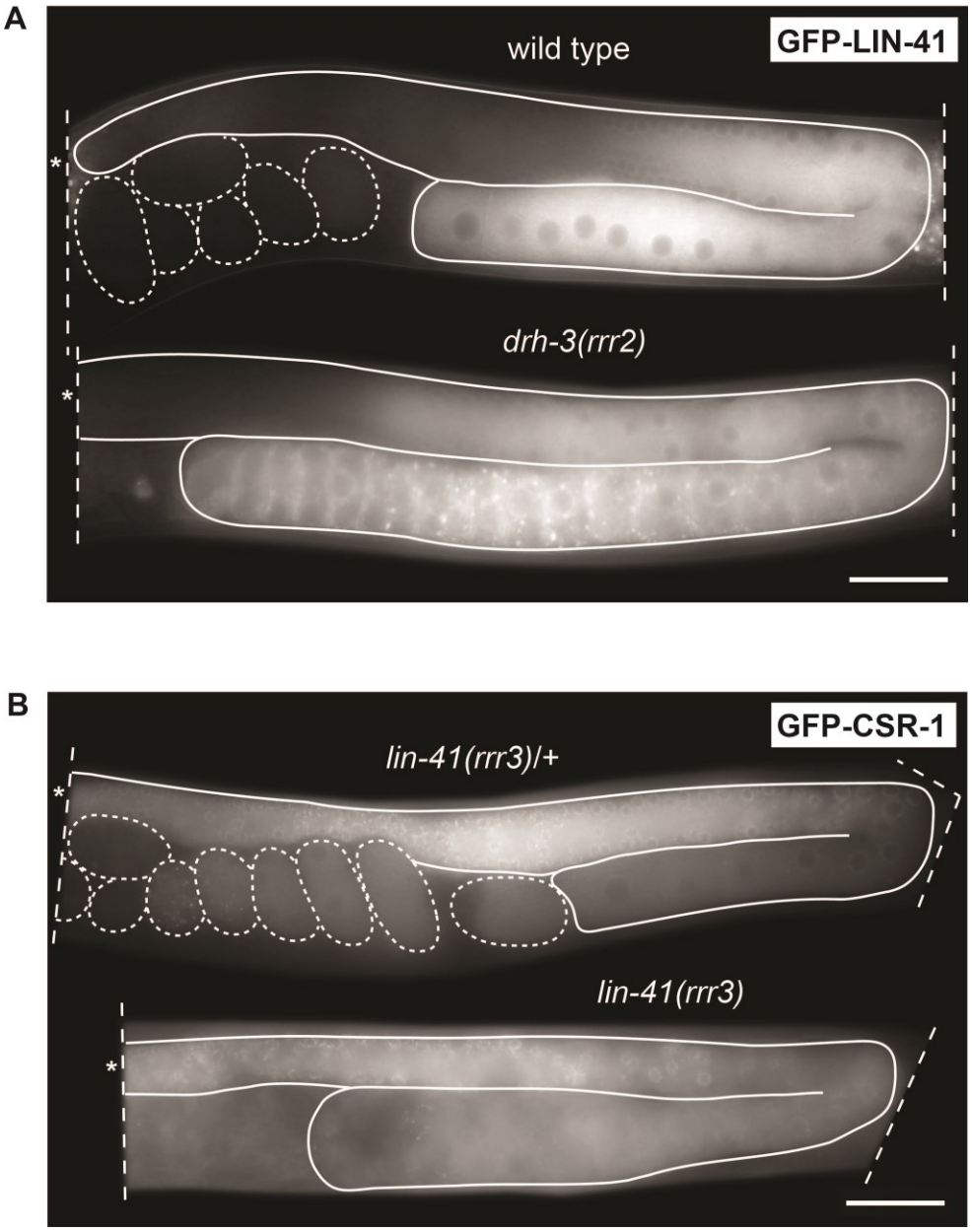


Figure 3.8: LIN-41 does not regulate CSR-1 expression, or *vice versa*

A. Fluorescent micrographs of live animals expressing a GFP-LIN-41 construct. Wild type animals and *drh-3(rrr2)* mutants show comparable expression of LIN-41 in the germline. Gonads are outlined with a continuous line, embryos with a dashed line and asterisks mark the distal end of the gonad. Scale bar: 40 µm. **B.** Fluorescent micrographs of live animals containing an endogenously GFP-tagged CSR-1. *Lin-41(rrr3)/+* heterozygote animals show cytoplasmic, perinuclear localization of CSR-1 in germ cells. *Lin-41(rrr3)* homozygous mutants show a similar CSR-1 expression and localization, except for the proximal germline where a teratoma is formed and thus, the germline-to-soma transition leads to a degradation of germline-specific components (Tocchini et al. 2014). Scale bar: 40 µm.

Another possibility for a common function of LIN-41 and CSR-1, is the requirement of LIN-41 in the CSR-1 pathway. To address if LIN-41 could contribute to the biogenesis or stability of CSR-1 22G RNAs, we sequenced small RNAs from *lin-41*-depleted animals. As controls served samples from *drh-3* and *ego-1*-depleted animals, where major changes in small RNA abundance are expected due to their function in 22G RNA biogenesis (Gu et al. 2009; Claycomb et al. 2009). Indeed, when we compared the small RNA reads detected in *drh-3* and *ego-1* knockdown animals compared to control animals, we could detect a depletion of endo-siRNAs targeting protein coding genes, which are expected to be comprised of mostly 22G RNAs (Figure 3.9). Another small RNA class, called 21U RNAs (piRNAs), was increased. A similar decrease in endo-siRNAs antisense to coding genes and an enrichment in piRNAs has been observed in small RNA samples from *drh-3* mutants before (Gu et al. 2009). The changes we saw in our knockdown animals are less dramatic compared to *drh-3* mutants, which is likely caused by a partial knockdown efficiency due to the role of DRH-3 and EGO-1 in exogenous RNAi. Together, the changes in small RNA classes in *drh-3* and *ego-1* knockdown animals showed that our small RNA sequencing results are of good quality and would allow us to detect changes in small RNAs, also in the *lin-41* knockdown animals. Depletion of LIN-41 resulted in a minor decrease in endo-siRNAs antisense to coding genes in replicate one, but not in replicate two. Together, *lin-41*-depletion did not result in consistent changes in small RNA classes (Figure 3.9). To have a more specific view on the effects on CSR-1 22G RNAs, we highlighted those small RNAs that have been identified to be associated with CSR-1 (Claycomb et al. 2009) in the small RNA scatter plots (Figure 3.10). In *drh-3* and *ego-1* knockdown animals, there is a decrease in CSR-1 22G RNAs detected in the second replicates only. This is likely due to a varying knockdown efficiency between the two replicates. In *lin-41*-depleted animals, CSR-1 22G RNA levels stayed the same. In parallel to small RNA sequencing, we sequenced total RNA from the same samples to correlate the changes in small RNAs with the effects on the corresponding mRNAs. *Drh-3* and *ego-1* knockdown did not have an effect on mRNA levels (Figure 3.11). However, as our RNA-seq data from *drh-3(rrr2)* mutants showed up- and down-regulated genes (Figure S3.2), we presume this is due to the partial knockdown of the proteins. Knockdown of *lin-41*, in contrast, showed many up- and down-regulated mRNAs compared to control. This is expected given the very strong somatic- and germline phenotypes upon loss of LIN-41 and revealed an efficient knockdown of *lin-41* in our samples. Unfortunately, the lack of changing mRNAs after *drh-3* and *ego-1* depletion did not allow us to identify commonly misregulated genes with *lin-41*. Also, when analyzing the mRNA levels of those few genes with changing small RNAs in *lin-41*, we did not see changes in mRNA levels. In total, although an efficient knockdown of *lin-41* was achieved, we could not detect

changes on small RNA levels in general, nor on CSR-1 22G RNA levels in particular. Therefore, it seems that LIN-41 does not play a role in CSR-1 22G RNA synthesis.

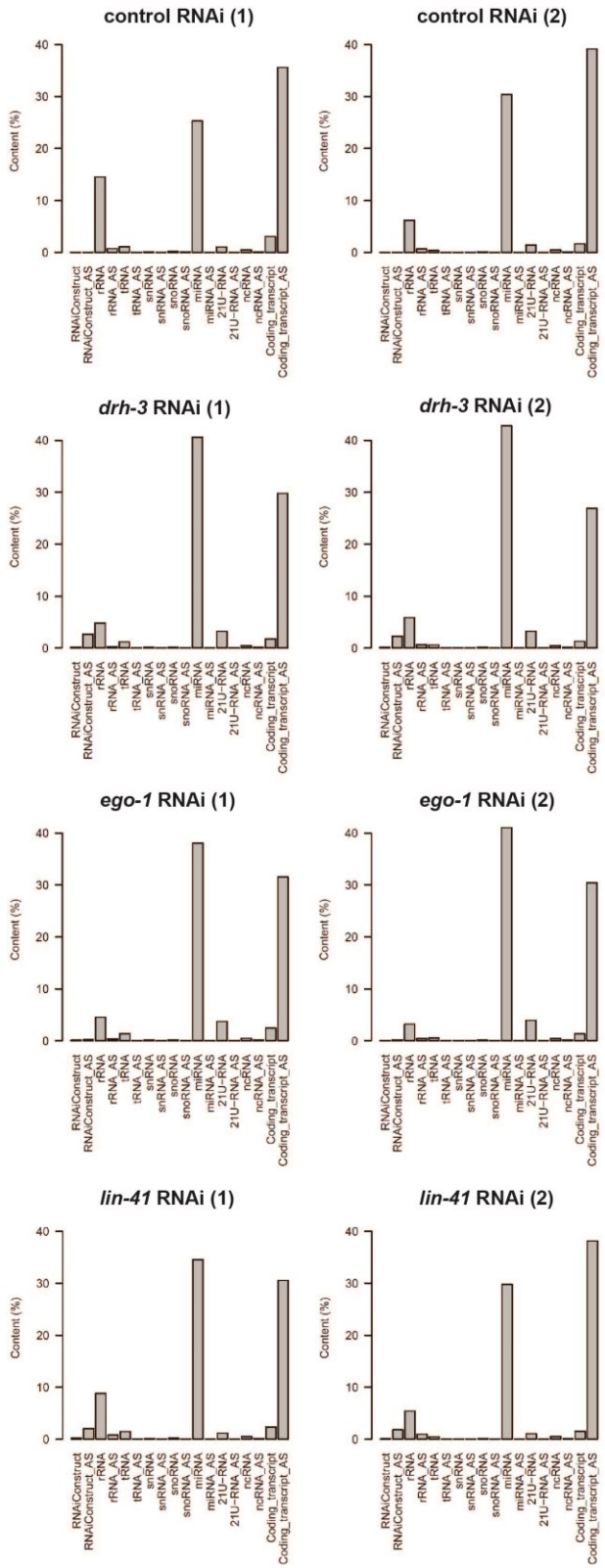


Figure 3.9: Small RNA populations detected in *drh-3-, ego-1-* and *lin-41*-depleted animals compared to control

The plots show the percentage of reads mapping to different small RNA classes, sequenced from control (empty vector), *drh-3*, *ego-1,* and *lin-41*-depleted animals. The samples were processed in duplicates.

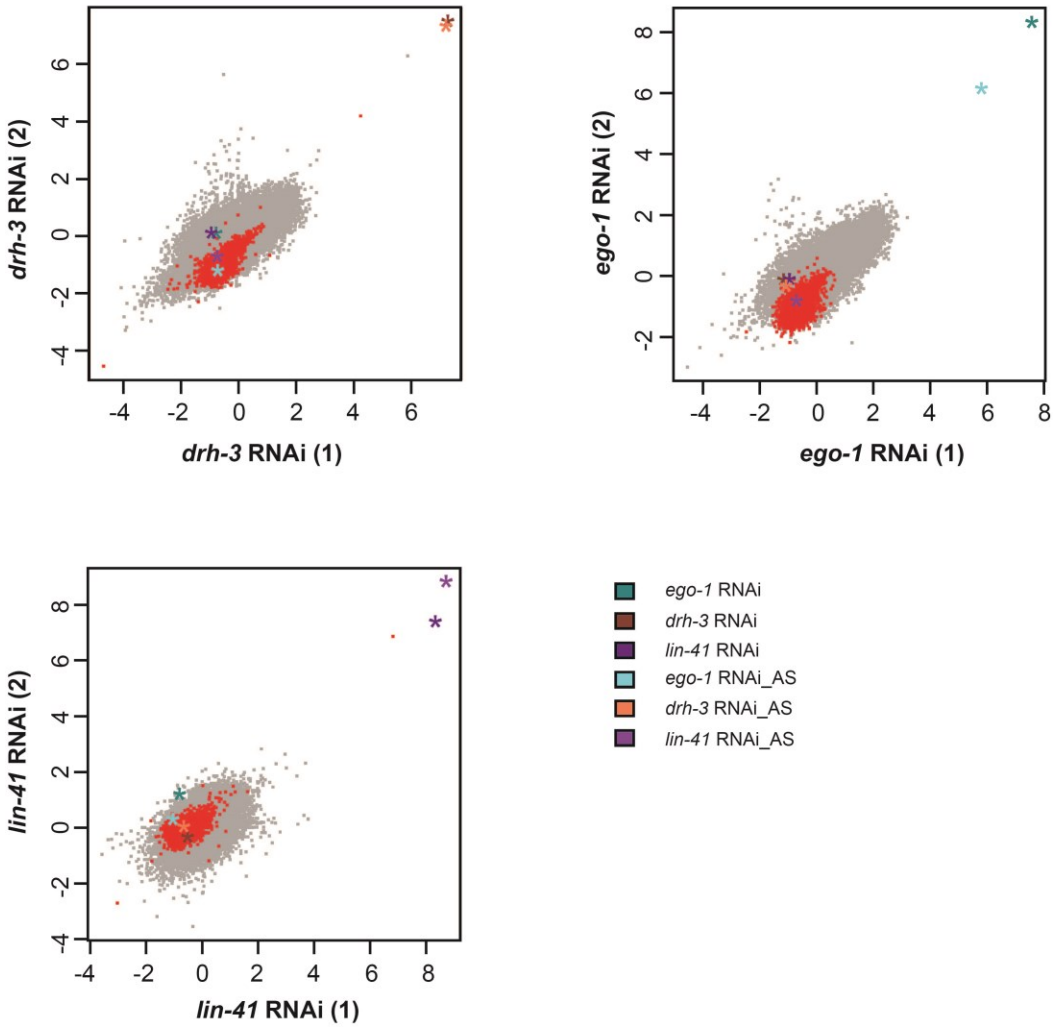


Figure 3.10: Small RNAs are not affected in *lin-41*-depleted animals

The scatter plots show expression changes in small RNAs (log_2) in *drh-3-, ego-1-,* and *lin-41*-depleted animals compared to control (empty vector RNAi). The samples were processed in duplicates. In red colored are CSR-1 22G RNAs, defined by the small RNAs associated with CSR-1 (Claycomb et al. 2009). Stars represent the small RNAs mapping to the respective RNAi constructs, as indicated in the legend. Changes in small RNA expression are detected in *drh-3* and *ego-1*-depleted animals. Small RNAs in *lin-41*-depleted animals show no major changes.

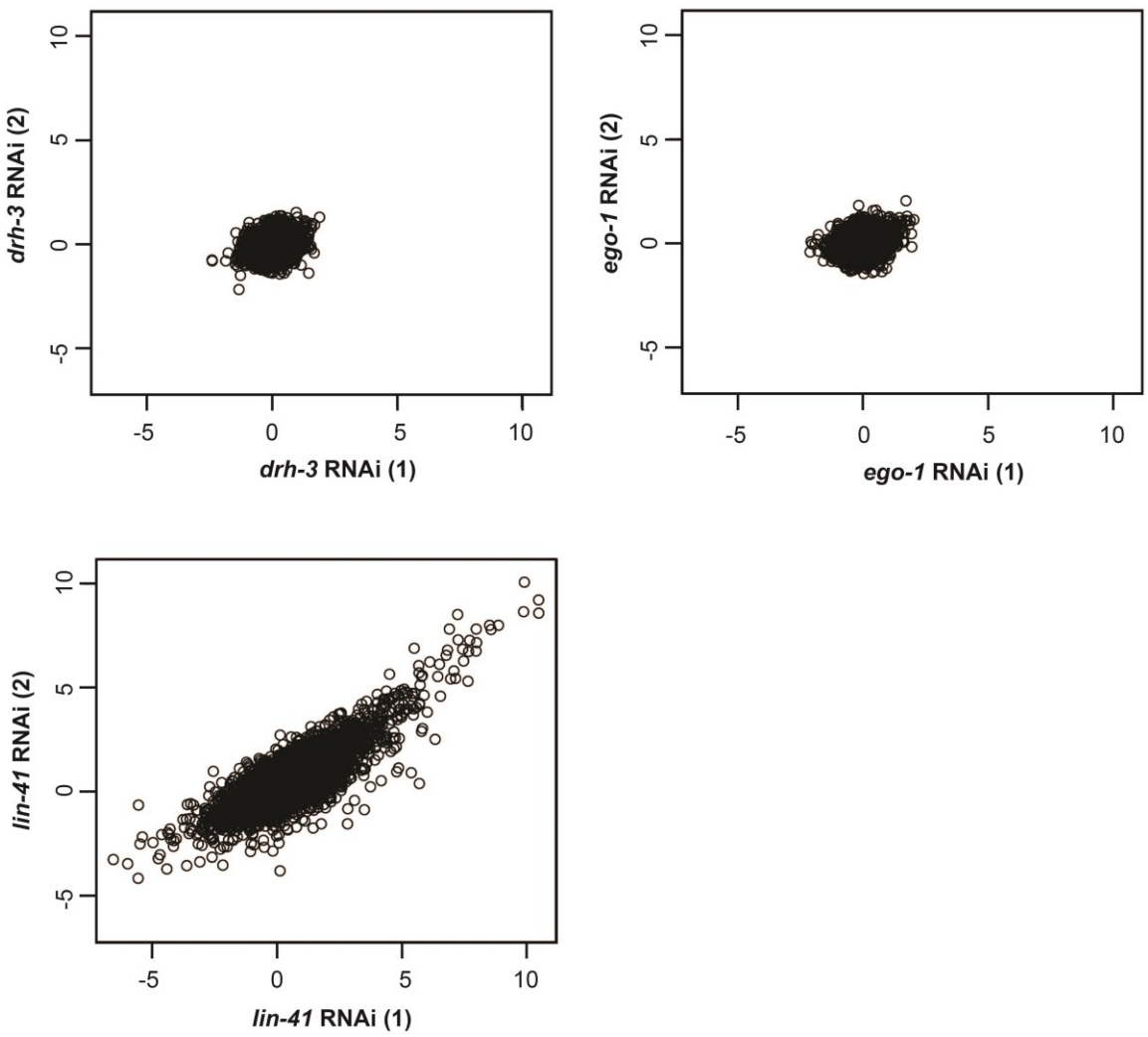
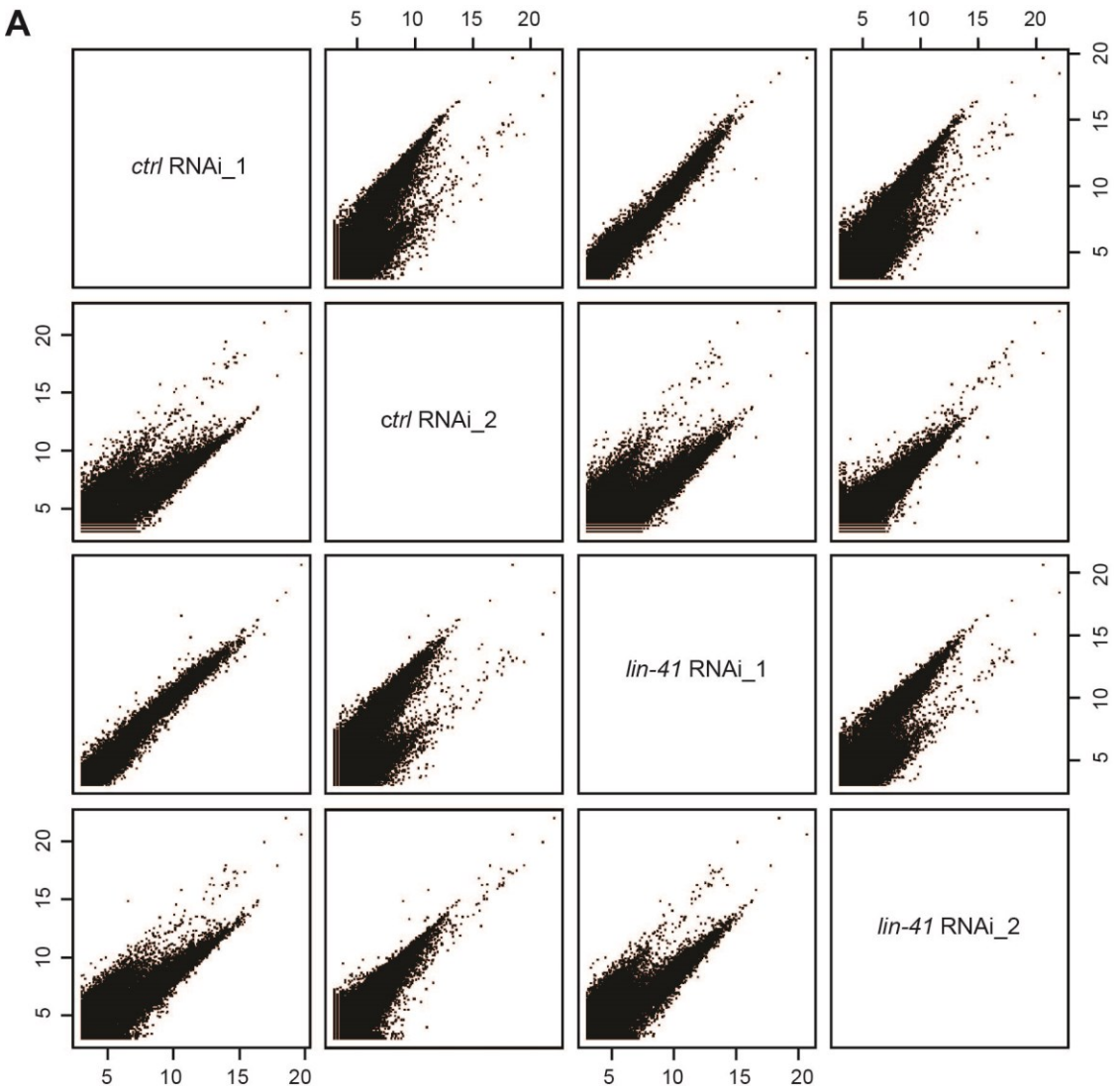


Figure 3.11: RNA sequencing of *drh-3, ego-1* and *lin-41*-depleted animals

The scatter plots show expression changes in mRNAs (log_2) in *drh-3-, ego-1-,* and *lin-41*-depleted animals compared to control (empty vector RNAi). The samples were processed in duplicates. *Drh-3* and *ego-1*-depleted animals show no changes in mRNA levels, probably due to a partial knockdown efficiency. *Lin-41*-depleted animals reveal big changes on the mRNA level, as expected from its strong soma- and germline-phenotypes, and indicates a successful knockdown.

CSR-1 binds to its targets in a 22G RNA-dependent manner. Thus, for a functional CSR-1 pathway, not only the biogenesis of 22G RNAs has to be intact, but also the loading of 22G RNAs to the Argonaute. To study a potential involvement of LIN-41 in the loading of 22G RNAs to CSR-1, we sequenced CSR-1-associated small RNAs from control and *lin-41*-depleted animals. The first replicate was collected from young adult animals and a second replicate from adult animals. The scatter plots showed strong changes in small RNA expression between the two replicates (Figure 3.12A). These differences can be likely explained by the sample harvesting from different stages. When comparing control animals to *lin-41*-depleted animals within one replicate, we did not see any changes in small RNA expression in young adults and only minor changes in adults. Moreover, when we compared the expression changes in between the two replicates, we could not detect any common changes (Figure 3.12B). Therefore, it seems that LIN-41 does not play a role in the loading of small RNAs to CSR-1.


A
ctrl RNAi_1
ctrl RNAi_2
lin-41 RNAi_1
lin-41 RNAi_2
5
10
15
20

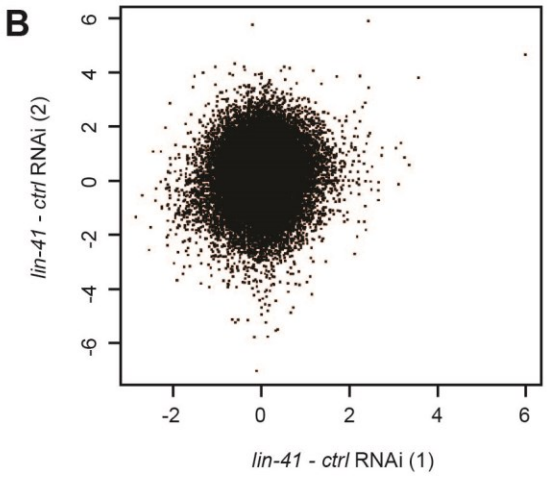

B
lin-41 - ctrl RNAi (2)
lin-41 - ctrl RNAi (1)
-6
-4
-2
0
2
4
6

Figure 3.12: Small RNA sequencing from CSR-1 complexes

A. The pairwise correlation plots show the small RNAs (log_2) detected in CSR-1 immunoprecipitation experiments from *lin-41* RNAi-treated animals compared to control (empty vector) RNAi-treated animals. Replicate 1 is collected from young adult worms, replicate 2 from adult worms. Strong changes in small RNA populations occur between the two replicates. However, small RNAs in *lin-41* RNAi-treated animals correlate well with control RNAi-treated animals in the same replicate, indicating that *lin-41* depletion does not have an effect on the loading of small RNAs into CSR-1 complexes. **B.** Comparison of the expression changes in between the two replicates did not reveal common changes.

3.2.2 Discussion

The discovery of a direct interaction between the two EGA repressors LIN-41 and CSR-1 *in vivo,* led us to speculate about a common function of the two proteins in inhibiting a precocious EGA in the germline. We could confirm a direct interaction of the proteins *in vitro* and map the interacting domains to the LIN-41 NHL domain and the CSR-1 MID-PIWI domain. The requirement of the LIN-41 NHL domain in EGA repression has been previously shown by introducing a few point mutations in the domain, which presumably abrogate its ability to bind mRNAs (Tocchini 2015). Also for the CSR-1 PIWI domain, we have shown an *in vivo* function in EGA repression. The DDH motif conferring the CSR-1 slicing activity is located in this domain and has been shown to be required to suppress a precocious EGA (Chapter 3.1). If the binding of the proteins *via* these domains is linked to their functions in EGA repression, is an interesting aspect to study in the future. To address the *in vivo* function of this interaction in EGA repression, it will be necessary to make point mutations that specifically interfere with the binding, but not with other functions of the two proteins. Our approach to identify the amino acid residues required for the binding by using peptide microarrays did not allow us to identify residues with high confidence to be tested *in vivo*. This is partially due to the limitations of the peptide microarray, which can only detect interactions mediated *via* the primary amino acid sequence of the peptides printed on the array, but does not take into account the structure of the proteins. Another reason is the lack of information of the CSR-1 structure. This worm Argonaute protein is not very well conserved to the human AGO2, for which structural information is available. Therefore, we could not get reliable information on the structure of a potential peptide hit identified in the array. Thus, structural information of CSR-1 is crucial to make informed point mutations *in vivo,* and especially the co-crystallization of the CSR-1 MID-PIWI domain with the LIN-41 NHL domain could depict the important residues for their binding.

In our attempts to identify a functional relevance of this interaction, we first tested if LIN-41 directly regulates the expression level or localization of CSR-1, or *vice versa*. However, LIN-41 expression in *drh-3* mutants and CSR-1 expression in *lin-41* mutants appeared unaffected. We then tested the idea of LIN-41 being involved in the CSR-1 small RNA pathway. We sequenced total small RNAs from *lin-41*-depleted animals to address a role of LIN-41 in small RNA biogenesis. In addition, we sequenced CSR-1 associated small RNAs in *lin-41*-depleted animals, to address a function of LIN-41 in the loading of small RNAs into CSR-1. However, neither total 22G RNAs, nor 22G RNAs loaded into CSR-1 were affected by LIN-41 depletion, suggesting that LIN-41 does not play a role in the CSR-1 small RNA pathway.

Another possibility how LIN-41 and CSR-1 could function together in EGA repression is *via* a role for CSR-1 in the regulation of LIN-41 target mRNAs. Although, we did not test this possibility yet, I would like to present here some ideas for such a regulation. The ideas are based on a published role for CSR-1 in the FBF-1 pathway and on our discovery concerning the function of CSR-1 in EGA repression. In general, the abundance of a protein can be influenced on the mRNA level by the stability of the mRNA that is encoding the protein or by the translation rate of the mRNA. To distinguish between a function in mRNA stability or translation, we could perform ribosome profiling together with mRNA sequencing on *csr-1*- and *lin-41*-depleted animals. Consistent mRNA levels, but decreased ribosomal occupancy argues for a regulation by translational repression, whereas decreased mRNA levels with similar ribosomal occupancy argue for target mRNA degradation. Based on the mechanistic functions that have been described for CSR-1, it could participate in either of the two mechanisms to regulate the abundance of LIN-41 targets. Concerning the role of CSR-1 in translational repression, it is reported that CSR-1 is required for target mRNA repression of another RNA binding protein, named FBF-1 (Pumilio), by interacting directly with FBF-1 and the translation elongation factor EFT-3 (eEF1A) (Friend et al. 2012). FBF-1 is expressed only in the mitotic region of the adult germline (Crittenden et al. 2002) and can therefore not function in developing oocytes. However, we can speculate that LIN-41 could function similarly to FBF-1, to guide CSR-1 to its target RNAs in developing oocytes. According to the function of CSR-1 in the FBF-1 pathway, CSR-1 could contribute to mRNA regulation via the inhibition of the eEF1A GTPase activity resulting in translational inhibition (Figure 3.13 A). However, it has to be mentioned that this function appears to be small RNA-independent, because the defects in translational repression have not been observed in mutants of the RdRP complex components DRH-3 and EGO-1. However, the role of CSR-1 in EGA repression is likely small RNA-dependent, as the derepression of embryonic genes is similarly observed after loss of DRH-3 and EGO-1. Thus, the mechanistic functions of CSR-1 in EGA repression and in the FBF-1 pathway might differ.

Alternatively, CSR-1 could function in the degradation of LIN-41 target mRNAs. One possibility is that CSR-1 directly functions in the degradation of the mRNAs via its conserved endonuclease (slicing) activity (Yigit et al. 2006), which we showed is important for CSR-1's role in EGA repression (Figure 3.13B). Another possibility would be the recruitment of factors involved in mRNA deadenylation or decapping, resulting in degradation *via* exonucleolytic decay (Figure 3.13C). To test these ideas, we first need to identify target mRNAs of LIN-41 in the germline. LIN-41 RNA-immunoprecipitation (RIP) experiments have been performed in the lab and identified potential interacting mRNAs of LIN-41. After validating a direct interaction for some germline mRNAs with LIN-41, we could use those to address the question how they might be co-regulated by the LIN-41-CSR-1 complex.

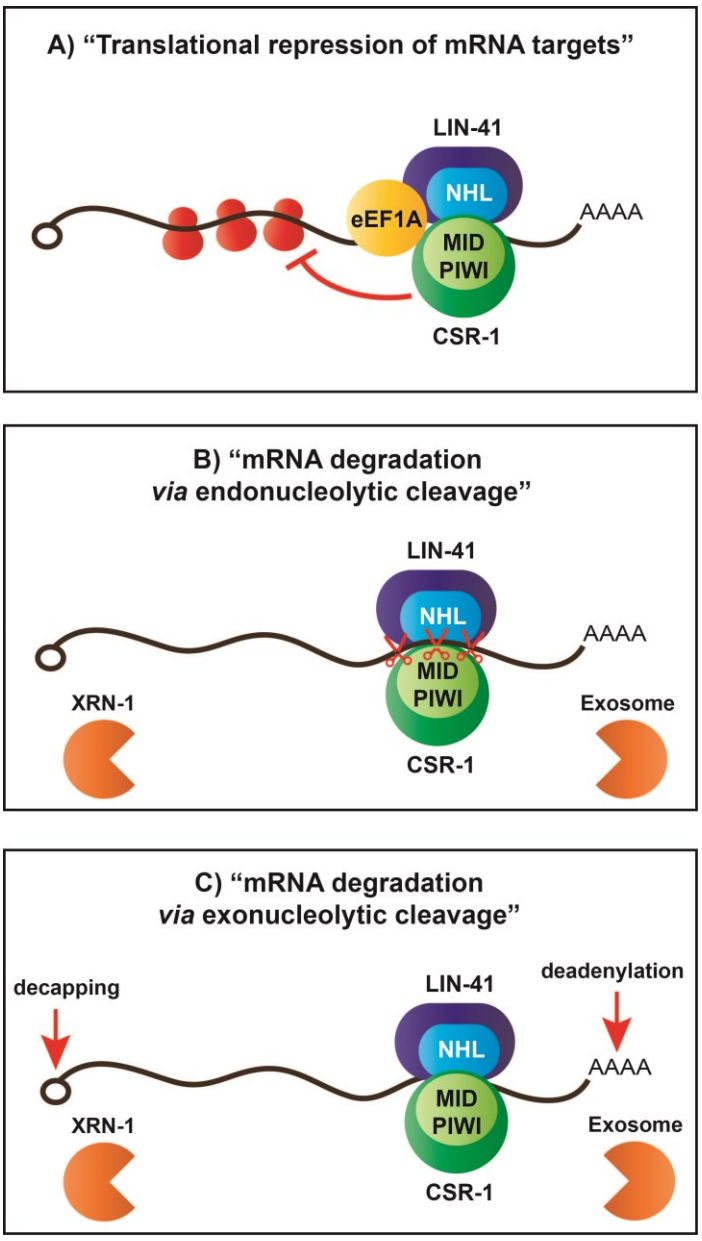


Figure 3.13: Different models for a co-regulatory role of CSR-1 and LIN-41

We propose that LIN-41 and CSR-1 can co-regulate target mRNAs in the germline, either A) by translational repression, where CSR-1 and LIN-41 could regulate eEF1A, similar to the function that has been reported for the CSR-1/FBF-1/eEF1A complex (Friend et al. 2012) B) by endonucleolytic cleavage of the target mRNAs mediated by the catalytic slicer activity of CSR-1 C) by recruiting additional regulators of mRNA stability, such as factors involved in deadenylation or decapping.

4 General Discussion

The reprogramming occurring during the OET is a complex process and we are only at the beginning of understanding the underlying mechanisms. Work from various model organisms has shown that key processes appear to be conserved during the OET. They include the silencing of transcription in oocytes, which is followed by a reactivation of transcription in early embryos (embryonic genome activation, EGA), and the occurrence of EGA in two waves. The mechanisms controlling these processes work at the transcriptional and post-transcriptional levels and may vary between organisms. Thus, one future challenge is to address the conservation of the identified players and mechanisms. The discovery of the pioneering transcription factor Zelda, and its promoter-binding motif in embryonic genes, enhanced our knowledge about the transcriptional regulation of embryonic genes in *Drosophila*. However, DNA-binding motifs specific to embryonic genes, or the transcription factors binding to them, have not been identified in other model organisms, which leaves the question about the conservation of this mechanism open. Moreover, while some of the identified players establish a global transcriptional competence in early embryos, the mechanisms mediating specificity of gene expression remain poorly understood. Furthermore, posttranscriptional regulation during the OET remains largely unexplored, although it is expected to play an important role during the transcriptionally silent period preceding EGA. Advances in the understanding of the post-transcriptional mechanisms contributing to reprogramming mostly come from *C. elegans.* There, the RNA-binding proteins GLD-1 and LIN-41 have been identified to contribute to EGA regulation. This thesis provides evidence for the importance of another type of post-transcriptional regulation to control EGA, *via* endogenous-small interfering RNAs of the CSR-1 Argonaute pathway. Importantly, we find that the catalytic activity of CSR-1 to degrade RNAs is required for this function. However, the precise molecular mechanism and the conservation of this functions remain to be explored. Once the individual regulatory mechanisms contributing to EGA control are understood, the next challenge will be to understand the interplay between those mechanisms. While some regulatory hierarchies have been understood, for example the regulatory axis in *C. elegans* embryos connecting the post-translational modification of OMA-1 and OMA-2 proteins to the cytoplasmic retention of the transcription factor TAF-4, the connections between other types of regulatory mechanisms remain to be determined. Dissecting the mechanisms underlying reprogramming during the OET will not only enhance the understanding of one of the most important developmental events but can also help in understanding and treating diseases, in which uncontrolled reprogramming can lead to teratomatous differentiation or undifferentiated tumors. Moreover, the identification of novel players controlling reprogramming is of

potential interest for the production of induced pluripotent stem (iPS) cells. The production of iPS cells still remains inefficient, with 0.01% - 6% efficiency by transcription factor overexpression (Pasque et al. 2011), and a huge effort goes into developing strategies to achieve a higher reprogramming efficiency. Basic research on reprogramming can contribute to this field by uncovering new molecular players and pathways, potentially with conserved human functions.

5 Annex

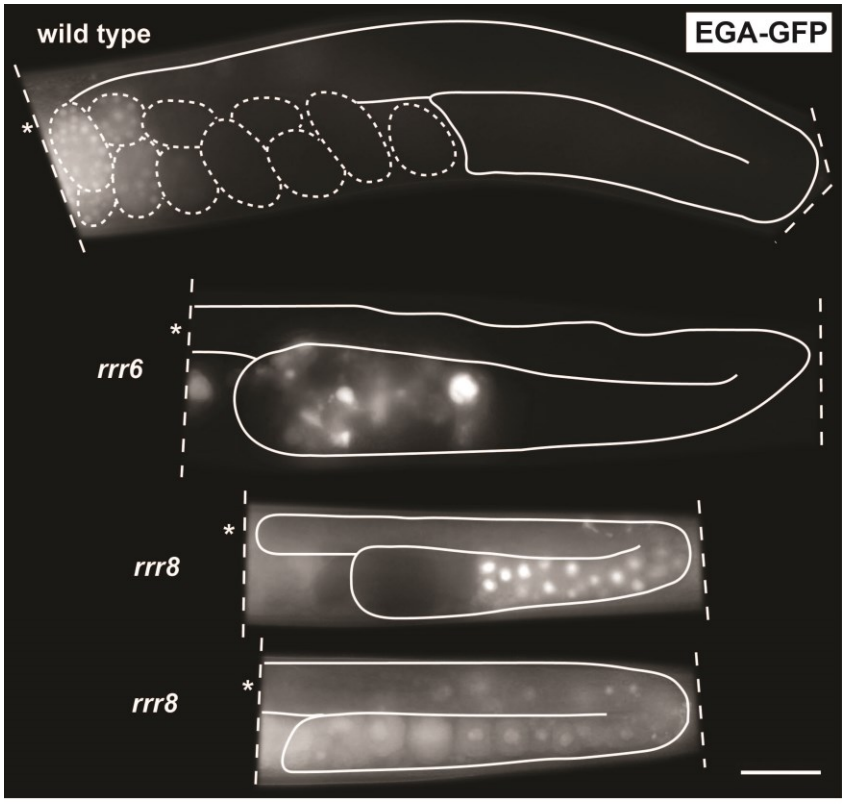
5.1 Additional mutants identified in the mutagenesis screen

5.1.1 Results

As previously described in chapter 3.1, I performed an EMS mutagenesis screen using an embryonic gene reporter to identify mutants that precociously express the reporter in the germ cells and hence may have mutations in EGA repressor genes. In addition to the small RNA pathway components *drh-3* and *ego-1*, described earlier, I identified two more potential EGA repressors. Furthermore, while performing the screen for EGA repressors, I also looked for EGA activators by screening for animals with embryos lacking the expression of the EGA reporter. I discovered one mutant showing such a phenotype. In this chapter, I will describe these additional mutants, with a preliminary description of their phenotypes and candidate mutations responsible for the defects in EGA regulation.

Mutant phenotypes - EGA repressors

The two potential EGA repressors, named *rrr6* and *rrr8* henceforth, showed, in addition to the precocious EGA reporter expression in germ cells, a fully penetrant sterile phenotype. However, they differ in their germline morphology. The *rrr6* mutant germlines appeared to first make some relatively normal oocytes, but these oocytes did not persist until the proximal part of the germline, where individual cells could not be distinguished anymore (Figure 5.1). Although this phenotype was not characterized in more detail, the disappearance of individual cell shapes and the patchy expression of the EGA reporter in the proximal germline, reminded of the teratoma phenotype observed in *lin-41* mutants. The *rrr8* mutant germlines showed varying phenotypes (Figure 5.1). Some mutant animals contained an overall smaller germline, which might be attributable to decreased germ cell proliferation, and major defects that did not allow the formation of oocytes. The very proximal germline of those animals did not contain any cells, leaving a vacuole-like space. Other *rrr8* mutant animals appeared to have a less affected germline, which allowed the formation of oocyte-like cells. However, these oocytes differed considerably from wild type oocytes in their size and overall morphology. In sum, the *rrr6* and *rrr8* mutant germlines are more strongly affected than the *drh-3* and *ego-1* germlines, suggesting further defects in germ cell development in these mutants, in addition to the precocious EGA.


wild type
EGA-GFP
rrr6
rrr8
rrr8

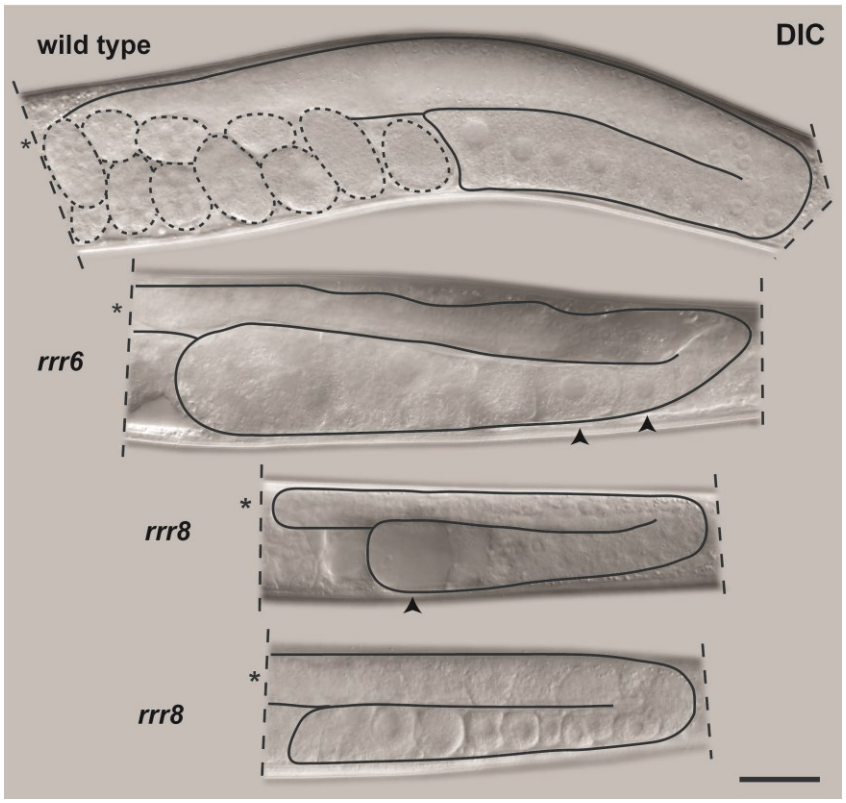

wild type
DIC
rrr6
rrr8
rrr8

Figure 5.1: *rrr6* and *rrr8* mutants express the EGA reporter precociously in germ cells

Fluorescent and DIC micrographs of live animals. Wild type animals express the EGA reporter specifically in embryos, whereas *rrr6* and *rrr8* mutant animals precociously express the reporter in germ cells. Gonads are outlined with a continuous line, embryos with a dashed line and asterisks mark the distal end of the gonad. Arrowheads in the DIC micrograph indicate oocytes in the *rrr6* mutant and a vacuolous germ cell region in the *rrr8* mutant. Scale bars: 40 µm.

Mutant phenotype - EGA activator

Early *C. elegans* embryos have been reported to continue their development without embryonic transcription, only with the maternally deposited gene products, until the 100-cell stage (Edgar et al. 1994). We therefore expected mutant embryos that lack an essential EGA activator to continue their development until several cell divisions have occurred. In our screen, we initially identified potential EGA activators based on the inability of the mutant embryos to express the EGA reporter. As this could also lead to the identification of mutant embryos with severe developmental defects leading to an early embryonic arrest, we looked at the development of the mutant embryos in more detail. Phenotypic analysis of one mutant identified in the screen, named *rrr7*, showed that the mutant embryos have undergone several cell divisions during embryonic development (Figure 5.2A). Furthermore, as expected for an EGA activator mutant, these animals were embryonic lethal. To determine if the cause of the phenotype was coming from either oocytes or sperms, we mated *rrr7* mutant hermaphrodites with wild type males and looked for a rescue of the phenotype. Wild type sperm could not rescue the phenotype, suggesting a maternal-effect embryonic lethality. *rrr7* mutant germlines, however, showed no obvious defects (data not shown). To characterize the embryonic defects in more detail we performed time-lapse microscopy. I observed much slower embryonic divisions in the *rrr7* mutant embryos compared to wild type (data not shown). This suggested that embryonic divisions in this mutant did not occur normally. Performing DAPI staining on the mutant embryos to visualize their DNA, revealed defects in chromosome segregation, indicated by the incomplete separation of the chromosomes, which led to the formation of so-called chromatin bridges (Figure 5.2B). Later stage embryos also showed differences in chromatin content, indicated by big size differences between the nuclei. In summary, although the *rrr7* mutant fulfilled our initial screening criteria for an EGA activator, further analysis of the embryonic phenotype revealed severe defects in embryonic divisions and chromatin content, suggesting that the lack of EGA reporter expression in those embryos is likely the consequence of those other defects.

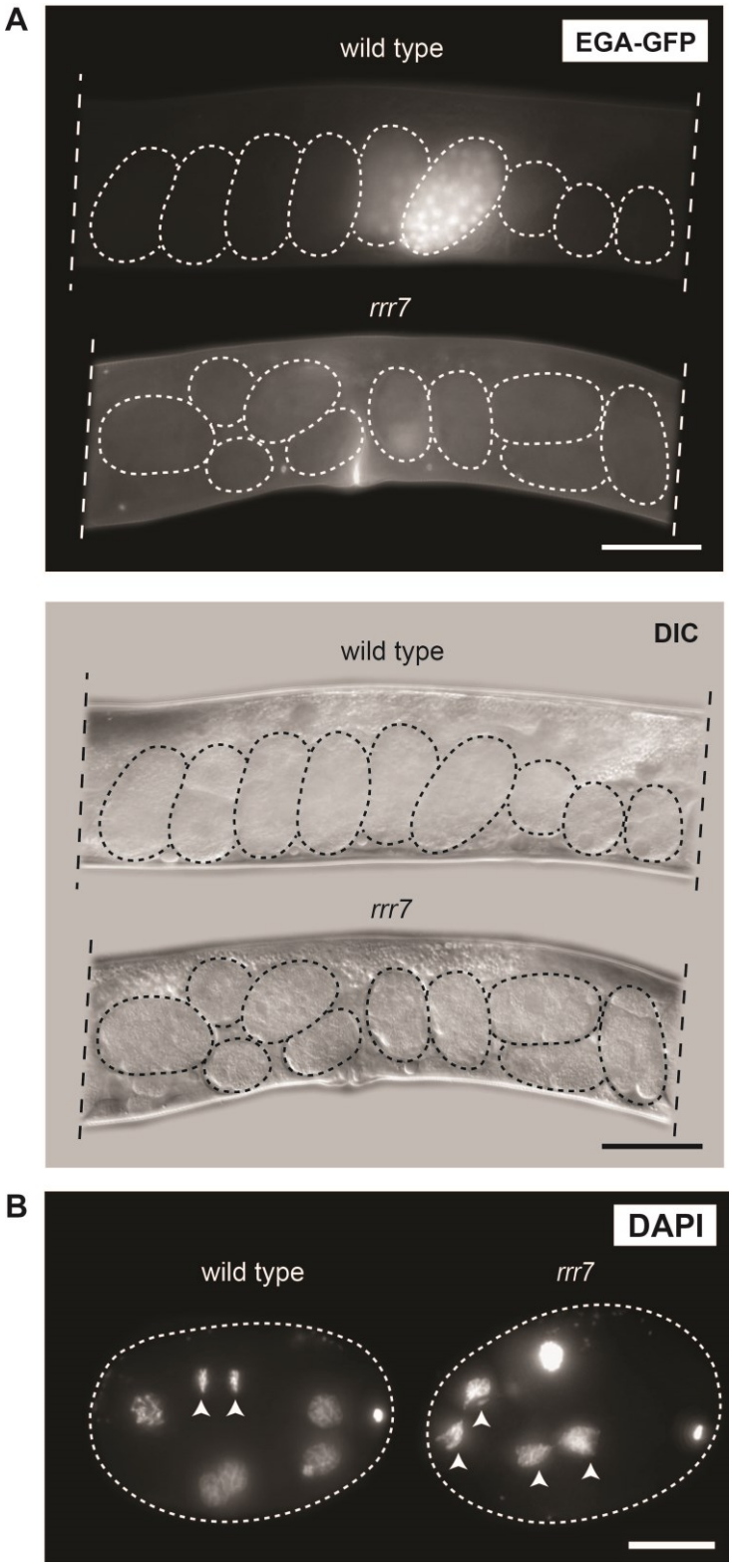

A
wild type
EGA-GFP
rrr7
DIC
wild type
rrr7
B
DAPI
wild type
rrr7

Figure 5.2: *rrr7* mutant embryos lack EGA reporter expression and show defects in chromosome segregation

A. Fluorescent and DIC micrographs of live wild type and *rrr7* mutant animals. Wild type embryos start expressing the EGA reporter, whereas *rrr7* mutant embryos lack reporter expression. Embryos are outlined with a dashed line. Scale bars: 40 µm. **B.** DAPI staining of wild type and *rrr7* mutant embryos to visualize DNA. *rrr7* mutant embryos show defects in chromosome segregation, indicated by the incomplete separation of chromatin between two nuclei (nuclei are indicated by arrowheads). Wild type embryos show completely separated chromatin (nuclei indicated by arrowheads). Embryos are outlined with a dashed line. Scale bar 15 µm.

Mapping of the mutants

To identify the phenotype-causing genetic mutations, we performed whole-genome-sequencing (WGS) and single nucleotide variant (SNV) mapping. The WGS data was filtered for SNVs of EMS-type (often GC to AT or AT to GC transitions (Sega 1984)) and SNVs that were not found in the parent strain. This SNV mapping revealed a cluster of mutations in the *rrr6* mutant on chromosome II (Figure 5.3). The phenotype-causing mutation in this mutant is not identified yet. There are several nonsynonymous mutations in protein-coding genes, which are likely candidates to cause the phenotype (Table 5.1).

The *rrr7* mutant contained only four SNVs, located on chromosomes III and X and on mitochondrial DNA (Figure 5.3). Balancing the mutation with an hT2 balancer covering regions of chromosomes I and III (McKim et al. 1993), suggested the mutation on chromosome III to be the phenotype-causing mutation. This mutation is a nonsynonymous mutation in the coding region of the *ran-2* gene (Table 5.1). RAN-2 is the *C. elegans* Ran GTPase-activating protein (RanGAP). RanGAPs enhance Ran activity by exchanging a guanosine diphosphate (GDP) to a guanosine triphosphate (GTP). Previous studies of *ran-2*-depleted embryos showed an important role for *ran-2* in mitotic spindle formation, chromosome segregation, nuclear envelope assembly, and co-migration of centrosomes with the male pronucleus (Askjaer et al. 2002). The reported phenotypes after *ran-2* depletion are very similar to the phenotypes we saw in our *rrr7* mutant embryos, further confirming the mapping of the phenotype-causing mutation to the *ran-2* gene.

The *rrr8* mutant contains four SNVs of EMS-type on chromosome I (Figure 5.3). The location of the mutation to chromosome I was additionally suggested by balancing the mutation with the hT2 balancer. To identify the phenotype-causing mutation among these, we performed RNAi-knockdown of the respective genes (Table 5.1) on the EGA reporter strain. The depletion of *usp-48* showed a similar phenotype to the *rrr8* mutant. Furthermore, the reported *usp-48* depletion phenotypes such as sterility and a protruding vulva phenotype also match the *rrr8* mutant phenotypes. Therefore, the

nonsynonymous mutation in the *usp-48* coding region is very likely causing the phenotype. *Usp-48* encodes a deubiquitinating enzyme (DUB) and belongs to the subclass of ubiquitin-specific proteases (USPs) (Papaevgeniou and Chondrogianni 2014). DUBs can remove monoubiquitin and polyubiquitin chains from proteins by cleaving the peptide or isopeptide bond between ubiquitin and the protein. Thereby DUBs can counteract the ubiquitination pathway and its canonical role in protein degradation. *Usp-48* has been identified as a synthetic multivulva (SynMuv) suppressor gene suggesting a role for *usp-48* in vulval cell-fate specification (Cui et al. 2006). If the function of *usp-48* in the ubiquitination pathway is important for this role remains unknown. As many SynMuv genes encode transcription factors and chromatin remodelers, *usp-48* has been suggested to function rather at the chromatin-level.

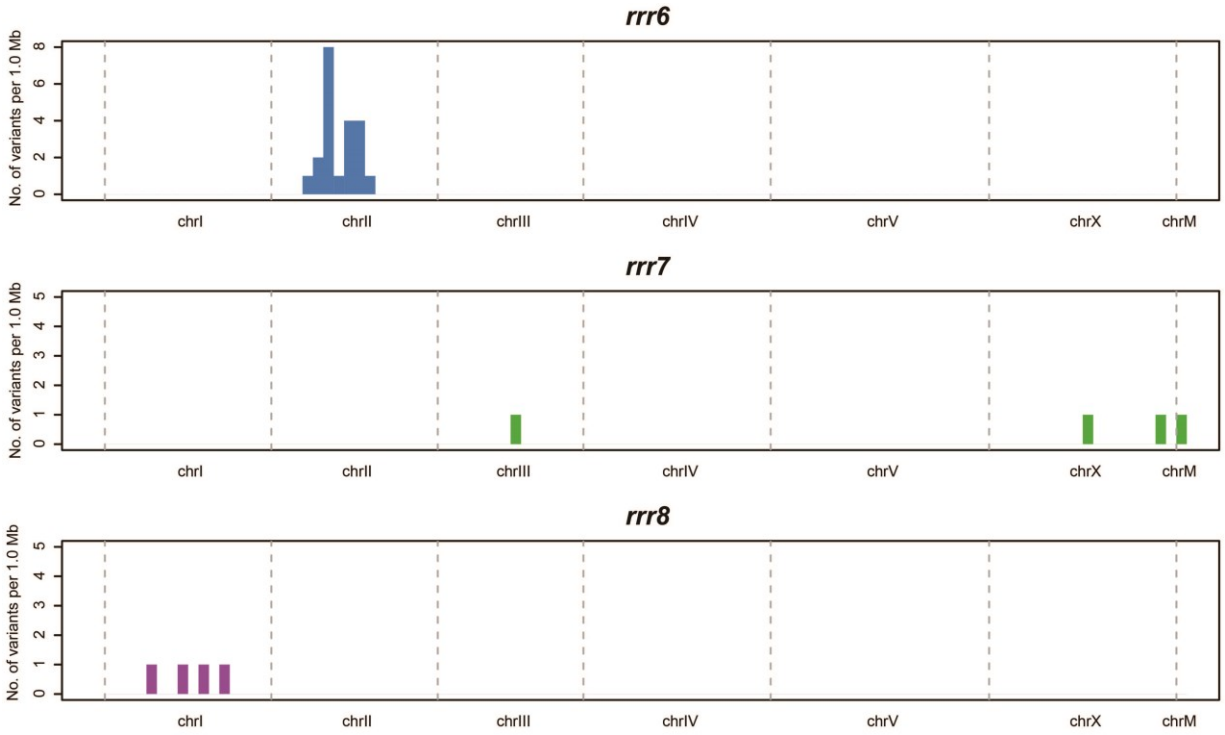


Figure 5.3: Single-nucleotide variants (SNVs) in the mutants induced by EMS-mutagenesis

After EMS-mutagenesis, all mutants were outcrossed against the wild type parental strain 5 to 8 times. Whole-genome-sequencing (WGS) revealed SNVs clustering on chromosome II for the *rrr6* mutant. The *rrr7* mutant contains SNVs on chromosomes III, X and on mitochondrial DNA. The *rrr8* mutant shows several SNVs on chromosome I. Numbers indicate chromosomes and “M” mitochondrial DNA. WGS data was filtered by "high quality SNVs (single nucleotide variants) of EMS-type not found in parent".

Strain	Chr.	Genomic position	Location	Gene ID	Gene name	Type of mutation
rrr6	II	3773054	coding	WBGene00020777	*T24E12.10*	synonymous
	II	4279503	intergenic	NA	NA	NA
	II	4944698	coding	WBGene00022765	*ZK546.14*	nonsynonymous
	II	5016097	intergenic	NA	NA	NA
	II	5129183	promoter	WBGene00001588	*F09C12.1*	NA
	II	5367127	intron promoter	WBGene00017284 WBGene00017292	*F09E5.5.* *F09E5.16*	NA
	II	5383519	coding	WBGene00004034	*F09E5.1 (pkc-3)*	nonsynonymous
	II	5510470	coding	WBGene00022673	*ZK177.8*	synonymous
	II	5512653	promoter intron	WBGene00022673	*ZK177.8*	NA
	II	5597991	coding	WBGene00015915	*C17G10.1*	nonsynonymous
	II	5804053	intergenic	NA	NA	NA
	II	6498749	promoter	WBGene00019904	*R05G9.2 (twk-49)*	NA
	II	7273112	promoter	WBGene00016275	NA	NA
	II	7417922	intergenic	NA	NA	NA
	II	7484264	promoter	WBGene00000280	*D1022.8 (cah-2)*	NA
	II	7854142	coding	WBGene00007014	*T09A5.6 (mdt-10)*	nonsynonymous
	II	8057481	promoter	WBGene00011512	*T05H10.8*	NA
	II	8224311	promoter	WBGene00000894	*M110.5 (dab-1)*	NA
	II	8351720	coding	WBGene00007746	*C26D10.6*	nonsynonymous
	II	8595698	coding	WBGene00009203	*F28C6.2 (aptf-3)*	synonymous
	II	9258152	intergenic	NA	NA	NA
rrr7	III	7934127	coding	WBGene00004303	*C29E4.3 (ran-2)*	nonsynonymous
	M	759	coding	WBGene00010958	*MTCE.4 (ndfl-4)*	nonsynonymous
	X	9863960	promoter promoter	WBGene00011077 WBGene00001784	*R07B1.5* *R07B1.4 (gst-36)*	NA
	X	16044352	promoter	WBGene00008104	*B0395.t1*	NA
rrr8	I	4855314	intron	WBGene00015303	*C01F4.2 (rga-6)*	NA
	I	7644331	promoter intron	WBGene00009141 WBGene00004508	*F26A3.2 (ncbp-2)* *F26A3.8 (rrf-1)*	NA
	I	9510334	coding	WBGene00009267	*F30A10.10 (usp-48)*	nonsynonymous
	I	11292726	coding	WBGene00013181	*Y53H1B.2*	nonsense

Table 5.1: Summary of the candidate mutations detected in the isolated mutants

The table summarizes the SNVs detected in the *rrr6, rrr7* and *rrr8* mutants. WGS data was filtered by "high quality SNVs (single nucleotide variants) of EMS-type not found in parent".

5.1.2 Discussion

The additional putative EGA repressors identified in the screen provide an opportunity to increase our understanding of the mechanisms that regulate EGA during *C. elegans* development. The phenotype-causing mutation in the *rrr6* mutant has not been identified yet, but the location has been mapped to chromosome II and likely candidates have been identified by SNV mapping. The EGA phenotype in the *rrr8* mutant is most likely caused by the mutation in *usp-48. Usp-48* encodes for an ubiquitin-specific protease, which has been suggested to function in chromatin remodeling in *C. elegans* (Cui et al. 2006). If the function in the ubiquitin pathway, in chromatin remodeling, or an uncharacterized function of *usp-48* contributes to EGA regulation remains to be elucidated. Previous studies in our lab identified important roles of post-transcriptional gene regulation via RNA-binding proteins in the regulation of EGA (Ciosk et al. 2006; Biedermann et al. 2009; Tocchini et al. 2014). The function of the CSR-1 pathway in EGA control involves a post-transcriptional regulation *via* the CSR-1 slicing activity (Chapter 3.1). If the additional EGA repressors presented here, are involved in post-transcriptional regulation and if they function together with the identified players, or function in parallel pathways to repress EGA in the germline, remains to be determined. Compared to the CSR-1 pathway mutants, they exhibit more severe germline phenotypes suggesting additional defects to the misregulation of EGA. The identification of downstream target genes involved in EGA repression in CSR-1 pathway mutants has been restricted by their defects in exogenous RNAi, which does not allow a sufficient knockdown of target genes in the mutant backgrounds. An advantage of the *rrr6* and *rrr8* mutants to study EGA repression could be their usefulness for the identification of such downstream target genes via RNAi knockdown, assuming they are not involved in the RNAi pathway.

Our screen for EGA activators did not yield a candidate that is directly involved in the transcriptional regulation of embryonic genes. *Ran-2* mutant embryos initially seemed to match the expected phenotype of an EGA activator mutant showing embryonic development until several cell divisions and not expressing the EGA reporter. However, further analysis of the mutant embryos and of the phenotypes that have been published for *ran-2*-depleted embryos revealed severe defects in cell division and chromatin morphology suggesting that the failure to induce EGA is most likely a secondary effect. With the screen we expected to identify factors specifically involved in the transcriptional activation of embryonic genes, for example transcription factors (TFs) or chromatin regulators. Such factors have been discovered in other model organism to contribute to EGA. Zelda, for example, is a maternally deposited TF needed to start the transcription of more than a hundred embryonic genes in *Drosophila melanogaster* (Liang et al. 2008).

Our screen-setup allowed us to identify maternally or paternally deposited players in EGA initiation, by observing hermaphrodites of the second generation after mutagen treatment, for embryos without EGA reporter expression. That we did not identify such players can have several reasons. One putative reason is that such embryonic gene-specific TFs and chromatin regulators do not exist in *C. elegans*. In addition to screening for EGA activators, we also tried to identify TFs using other approaches, for example by analyzing embryonic gene promoters for TF binding motifs and depleting putative TFs associating with these motifs by RNAi. This approach also did not result in the identification of a TF needed to start embryonic transcription. A possible scenario how EGA could be controlled without embryonic gene-specific TFs, is via a global regulation of transcription, for example by controlling Pol II complex assembly or controlling Pol II activity. However, players required for transcription in general could not be discovered in our screen, as it allowed only the discovery of maternal- or paternal-effect embryonic lethal genes, but not genes directly required for embryonic development, as expected for basic transcription components. A mechanism that globally controls the transcription in *C. elegans* embryos has been discovered (Guven-Ozkan et al. 2008). Sequestration of TAF-4, a TFIID subunit, in the cytoplasm of one- and two-cell embryos inhibits the assembly of a functional Pol II complex and thereby transcription. The release of TAF-4 into the nucleus in the four-cell embryo, when EGA normally occurs, allows transcription to start. Although such a global regulation of transcription does not explain the specific transcription of embryonic genes, there are other mechanisms that could contribute to a gene-specific regulation. For example, specific features of embryonic genes, such as their short length as compared to all genes (Tocchini 2015), could contribute to their differential expression. The idea of a differential gene expression based on gene length follows the transcript abortion model, where during the short cell cycles of early embryos only the transcription of short genes is possible, whereas the transcription of longer genes is interrupted by the cell cycle, leading to the abortion of transcripts (Tadros and Lipshitz 2009). Another reason why we could not identify an EGA activator in the screen might be the redundancy of genes involved to initiate EGA. As EGA is very crucial for the development of the animal, redundant players or mechanisms could have evolved to ensure a successful embryonic development, even if one component is affected. A third reason for the lacking discovery of an EGA activator in our screen could be also due to screening limitations. However, this seems rather unlikely, because we covered the *C. elegans* genome more than once and our screen discovered several EGA repressors at the same time, which even led to the identification of several alleles of the same gene.

6 Material and Methods

The following Material and Methods have been used for chapters 3.1 and 5.1

Nematode culture, mutants and transgenic lines

Animals were maintained using standard procedures and were grown at 20°C, unless stated otherwise. Alleles and transgenic lines are separately listed at the end of the Material and Methods chapter.

Mutagenesis and whole-genome-sequencing (WGS)

EMS mutagenesis was performed as previously described (Brenner 1974) using a strain carrying the EGA-GFP reporter and a transgene to visualize P granules. Mapping of the mutants was performed as described in (Zuryn et al. 2010). Before WGS, each mutant was outcrossed 4-8 times to the unmutagenized parent strain. Genomic DNA was isolated using Gentra Puregene Tissue Kit 4 g (Qiagen). DNA libraries were made from 50 ng of genomic DNA using the Nextera DNA kit from Illumina. Sequencing was performed using Hi Seq 2000 from Illumina.

Processing of sequence data and detection of sequence variants

Sequence reads were aligned to the May 2008 C. elegans assembly (obtained from http://hgdownload.soe.ucsc.edu/goldenPath/ce6/chromosomes/) using ‘‘bwa’’ (Li and Durbin 2009); version 0.7.4) with default parameters, but only retaining single-hit alignments (‘‘bwa samse -n 1’’ for single reads and “bwa sampe -a 1000 -o 1000 -n 1” for paired end reads and selecting alignments with ‘‘X0:i:1’’). The resulting alignments were converted to BAM format, sorted and indexed using ‘‘samtools’’(Li et al. 2009); version 0.1.19). In order to quantify contamination by Escherichia coli, reads were similarly aligned to a collection of E. coli genomes (NCBI accession numbers NC_008253, NC_008563, NC_010468, NC_004431, NC_009801, NC_009800, NC_002655, NC_002695, NC_010498, NC_007946, NC_010473, NC_000913 and AC_000091), which typically resulted in less than 1% aligned reads. Potential PCR duplicates were identified and removed using Picard (version 1.92, http://broadinstitute.github.io/picard/), reducing the number of reads to 33% to 52% in single read samples, and to 93% in the paired-end read sample. Sequence variants were identified using GATK (DePristo et al. 2011) (version 2.5.2) following recommended “best practice variant detection”: Initial alignments were first corrected by indel realignment and base quality score recalibration, followed by SNP

and INDEL discovery and genotyping using “UnifiedGenotyper” for each individual strain using standard hard filtering parameters, resulting in a total of five to six thousand sequence variations in each strain compared to the reference genome. Finally, the number of high quality (score >=500) single nucleotide substitutions of EMS-type (G/C>A/T transitions (Drake and Baltz 1976)) not found in the parent strain (typically less than 1% of the total number of variants per strain) were counted in sequential windows of 1 Mb to identify regions of increased variant density.

Reverse transcription and quantitative PCR on dissected gonads

RNA was isolated from 50 gonads of 1-day-old (after the L4-to-adult molt) animals using the Picopure RNA Isolation Kit (Arcturus). 3 independent biological replicates were collected. cDNAs were generated using the QuantiTect Reverse Transcription Kit (Qiagen). Real-time PCR was performed in duplicates using Absolute QPCR SYBR green ROX mix (AbGene) on an ABI PRISM 7700 system or an StepOnePlus system (both Applied Biosystems). qPCR reactions were performed as described previously (Biedermann et al. 2009). At least one primer in each pair is specific for an exon-exon junction (see primer sequences at the end of the chapter). Standard curves for every primer pair were generated using a serial dilution of cDNA from embryos and were used to determine the amount of each transcript in the gonads. All technical duplicate values were first averaged, and then values for a particular transcript were normalized to the mean of the wild type. Error bars show the SEM for the three biological replicates.

EGA reporter assay

Temperature-sensitive mutants used for the EGA reporter assay were grown at 15°C. Mutant animals were synchronized by bleaching and put as L1 larvae on OP50 plates to 25°C until scored for EGA-GFP expression as adults.

Immunostaining

Immunostaining against SPD-2 (“969LA”, 1:800) was performed as previously described (Lin et al. 1998). A Zeiss Axio Imager Z1 microscope equipped with an Axiocam MRm REV 2 CCD camera was used for capturing pictures. Images were processed with Adobe Photoshop and Adobe Illustrator.

Protein Extraction and Western Blot Analysis

Synchronized adult worms were harvested for protein extraction as described before (Biedermann et al. 2009). Proteins were resolved by SDS-PAGE on 4-12% Bis-Tris Protein Gels (NuPAGE Novex) and transferred using the Trans-Blot Turbo Transfer System (BIO RAD). Membranes were blocked with 4% milk in PBST and incubated with the primary antibodies α-DRH-3 (Aoki et al. 2007), α-ACT-1 (MAB1501, Millipore), α-H2B (ab1790, Abcam) or α-PAB-1 (Scheckel et al. 2012) in blocking buffer at 4°C overnight. Membranes were washed three times in PBST and incubated with the secondary (HRP-conjugated) antibody (GE Healthcare) in blocking buffer at RT for 1 hour and again washed three times with PBST. Chemiluminescence was performed using Pierce ECL Western Blotting Substrate (Thermo Scientific).

I thank Hiroaki Tabara for sharing the DRH-3 antibody.

RNA interference

For RNAi, animals were fed with bacteria expressing dsRNAs (targeting *cdl-1* from the Ahringer library, or an "empty" vector as a negative control) from L1 stage. Animals were grown at 20°C and examined 65 hours post-L1 stage for expression of the EGA reporter in germ cells.

Selection of genes activated in the early embryo

We downloaded microarray expression data (Baugh et al. 2003) for the samples (GSM39513-GSM39519,GSM39543,GSM39522-GSM39526,GSM39530,GSM39531) from GEO (www.ncbi.nlm.nih.gov/geo/), representing the 4-cell (3x), 8-cell (4x), 15-cell (4x) and 26-cell (4x) stages. The data was normalized using the function justGCRMA from the Bioconductor package gcrma. We confirmed the high degree of consistency among the replicates and then averaged those to obtain one expression profile per stage. Early embryonic genes were defined as genes, which showed a change in expression of at least 2-fold from the 4-cell to the 8-cell or from the 8-cell to the 15-cell stages and no expression at the 4-cell stage (expression < 3.75). The data from the 26-cell stages was not used for selecting genes, as there was only little change in expression when comparing it to the 15 cell stage.

RNA sequencing and data analysis

For total RNA sequencing, the samples were prepared using the ScriptSeq v2 RNA-Seq Library Preparation Kit (Epicentre) and then sequenced on an Illumina HiSeq 2500. The total RNA sequencing data was analyzed as previously described (Hendriks et al. 2014).

The following Material and Methods have been used for chapter 3.2

Cell-free protein synthesis and co-immunoprecipitation of proteins

N-terminally tagged coding sequences of the respective proteins were cloned in the pCS2+ plasmid under the control of an SP6 promoter. Proteins were expressed from plasmids using the TNT SP6 High-Yield Wheat Germ Protein Expression System from Promega corresponding to manufacturer's protocol using 4-10 µg of plasmid in a 50 µl reaction for 2 h at 25°C. For protein binding, similar amounts of proteins were taken and allowed to bind for 2 h on ice in extraction buffer (50 mM HEPES, 100 mM KOAc, 5 mM MgAc, 0.1% Triton X-100, 10% Glycerol (w/v), 20 mM β-glycerophosphate) with protease inhibitors. Antibody (mouse α-Myc-Tag (9B11)) was allowed to bind for 1 h at 4°C, rotating. Immunoprecipitation was performed using magnetic dynabeads protein G from Novex for 3 hours at 4°C rotating. Washing was done twice with EB (+300mM NaCl), followed by one wash in EB. Elution was performed with 1x NuPAGE LDS sample buffer from Invitrogen. Samples were heated for 10 min at 70°C. Proteins were resolved by SDS-PAGE on 4-12% Bis-Tris Protein Gels (NuPAGE Novex) and transferred using the Trans-Blot Turbo Transfer System (BIO RAD). Membranes were blocked with 4% milk in PBST and incubated with primary antibody (mouse α-Flag (M2) 1:1000 or mouse α-Myc-Tag (9B11) 1:1000) in blocking buffer at 4°C overnight. Membranes were washed three times in PBST and incubated with HRP-conjugated α-mouse antibody (1:7500) in blocking buffer at RT for 2 hours and again washed three times with PBST. Chemiluminescence was performed using Pierce ECL Western Blotting Substrate.

Peptide Microarray

The protein binding studies using peptide microarrays were performed by JPT Peptide Technologies, Berlin, Germany. CSR-1 peptide sequences of 18 amino acids length were synthesized and immobilized on the microarray slide in the following order: CSR-1 isoform B (peptides 1-96), CSR-1 MID-PIWI domain, amino acid residues 557-993 in duplicates (peptides 97- 144 and peptides 145-192). The *C. elegans* LIN-41 FIL-NHL domain was expressed in insect cells and labeled using DyLight Microscale Antibody Labeling

Kit (Thermo Scientific 84536). The labeled protein was diluted to concentrations of 1 µg/mL or 0.2 µg/mL and incubated with the peptide microarrays. After washing and drying, the microarray slides were scanned using a high-resolution laser scanner to obtain fluorescence intensity profiles. Images were quantified and the MMC2-values of the mean pixel fluorescence were calculated for each peptide. The results were visualized in a heatmap diagram.

Small RNA cloning and sequencing

22G RNA isolation from total RNA and 22G RNA library formation were essentially performed as described in (Gu et al. 2011). In more detail the experiment was performed as follows: Wild type (N2) animals were grown on control (empty vector)/*drh-3*/*ego-1* or *lin-41* RNAi plates form L1 stage. The animals were harvested as young adults. Two independent biological replicates (prepared on different days) were prepared. RNA was isolated with TRIzol reagent. Small RNA purification was performed by loading 5 µg total RNA per sample on a 15% PAGE/7M Urea gel. Using a small RNA ladder as size marker, the small RNAs in between 20 and 25 nucleotides were cut out from the gel. The RNA was eluted in elution buffer (20mM Tris–HCl; 300mM NaCl; 5mM EDTA; 0.1% SDS) by smashing the gel with a plastic pestle followed by an incubation overnight at 4°C. The next day, the RNA was spun down using corning filter tubes and precipitated by adding 3 volumes of ethanol and incubating at -80°C for 1 hour. After washing the pellet with 80% ethanol and air drying the pellet, it was dissolved directly in 3' linker ligation mix (1 µl 10 µM 3'adapter oligo, 1 µl 10x T4 RNA Ligase Buffer, 2 µl 100% DMSO, 0.5 µl RNasin, 1.5 µl T4 RNA Ligase 2 truncated K227Q, 4 µl RNA). Ligation was performed overnight at 4°C. The 5' triphosphate of the 22G RNAs was removed by treating the samples with the *C. elegans* phosphatase PIR-1 for 30 minutes at room temperature. This treatment left a 5' monophosphate which can be ligated to the 5'linker. Before 5' linker ligation RT primer annealing was performed. The 5' linker (1 µl 10 µM 5'adapter oligo, 1.5 µl 10mM ATP, 1.5 µl T4 RNA ligase 1) was ligated at 20°C for 1 hour, followed by 37°C for 30 minutes. cDNA synthesis was performed with the Superscript III reverse transcriptase system (Invitrogen). To determine optimal cycle conditions for library amplification, pilot PCRs with increasing numbers of cycles (12-24) were performed. Large scale PCRs with the determined cycle number were loaded on a polyacrylamide gel and the amplified library was excised from the gel. The elution was performed as previously described and the DNA was precipitated with 3 volumes of ethanol and 1/10 the volume of sodium acetate overnight. The samples were sequenced on an Illumina HiSeq 2500.

Small RNA cloning and sequencing from CSR-1 complexes

FLAG::CSR-1 animals were grown on control (empty vector) or *lin-41* RNAi plates from L1 stage. Worms were harvested as young adults (replicate 1) or adult worms (replicate 2). Protein extraction was performed as described before (Biedermann et al. 2009). For the IPs, 11 mg protein were used for replicate 1 and 15 mg protein for replicate 2. IPs were performed using anti-FLAG M2 magnetic beads (Sigma) for 3 hours at 4°C rotating. After washing, the elution from beads was done with 125 µl elution buffer (50mM Tris; 5mM EDTA; 10mM DTT; 1% SDS) at 37°C for 15 minutes. For RNA extraction, 100 µl of the eluate were used. The remaining 25 µl were used for western blot analysis. RNA extraction was performed with the Picopure RNA Isolation Kit (Arcturus). Small RNA purification and library formation were done as previously described.

Processing of the small RNA sequencing data

The 3' adaptor (TGGAATTCTCGGGTGCCAAGG) was removed from the reads using the function preprocessReads from within the R package QuasR (version1.12.0, default parameters). Mapping of the short fragments to the *C. elegans* genome (ce10) was performed using bowtie (Langmead et al. 2009). We performed two runs of alignments, one allowing only for uniquely mapping reads (UM) and one allowing for unique, as well as multi mappers (MM). The commands used to perform the alignments were qAlign("samples.txt", "BSgenome.Celegans.UCSC.ce10") and qAlign("samples.txt", "BSgenome.Celegans:UCSC.ce10", maxHits=100), respectively. In the case of a multi mapping read, one alignment was selected randomly. We used the MM alignments to quantify the expression of broad annotation categories (e.g. rRNA, tRNA, snRNA) and the UM alignments to quantify at the single locus level. For each annotated feature two copies were created, one sense and one antisense, by flipping the respective strand. The parts of the genes covered directly by the used RNAi constructs were excluded from the quantification of the respective genes. Read counting was then performed using the command qCount (proj,exons,orientation="same"). Quantification of the broad annotation categories was performed by assigning each read hierarchically to the annotation classes in the following order: RNAiConstruct, rRNA, tRNA, snRNA, snoRNA, miRNA, 21U-RNA, ncRNA, coding_transcript (using the function qCount).

Total RNA sequencing

The total RNA samples are identical to those that we used for small RNA cloning, see details above. For total RNA sequencing, the samples were prepared using ScriptSeq v2 RNA-Seq Library Preparation Kit (Epicentre) and then sequenced on an Illumina HiSeq 2500.

Processing of the RNA-sequencing data

All the RNA-seq data were mapped to the *C. elegans* genome (ce10) with the R package QuasR (version 1.12.0) (Gaidatzis et al. 2015) using the included spliced alignment algorithm SpliceMap (Au et al. 2010). The command used to perform the alignments was "proj <- qAlign("samples.txt", "BSgenome.Celegans.UCSC.ce10", splicedAlignment=TRUE)". Gene expression was quantified by counting the number of reads that started within any of the exons belonging to a particular gene (WormBase, WS220). The command used to create the count table was qCount(proj,exons,orientation="same"). To compensate for differences in the read depths of the various libraries, we divided each sample by the total number of reads and multiplied by the average library size. Log2 expression levels were calculated after adding a pseudocount of 8 (y=log2(x+8). This was done to minimize the large differences in expression that would otherwise be caused by genes with small numbers of counts.

Alleles and transgenic lines used in this study

CGC or Ciosk lab number	**Genotype**
1284	*rrrSi199[Pvet-4::NLS:gfp:gfp::vet-4 3'UTR; unc-119(+)] (II); rrrSi198[Pvet-4::NLS:gfp:gfp::vet-4 3'UTR; unc-119(+)] (IV)*
1270	*rrrSi199 (II); rrrSi198 (IV); zuIs252[nmy-2::PGL-1::RFP, unc-119(ed3)]; glo-1 (zu391) (X)*
1350	*drh-3(rrr2) (I)/hT2 (I;III); rrrSi199 (II); rrrSi198 (IV); glo-1(zu391) (X)*
1431	*drh-3(rrr5) (I)/hT2 (I;III); rrrSi199 (II); rrrSi198 (IV); zuIs252; glo-1(zu391) (X)*
1600	*ego-1(rrr9) (I)/hT2 (I;III); rrrSi199 (II); rrrSi198 (IV); zuIs252; glo-1(zu391) (X)*
JK3025	*gld-1(q485) (I)/hT2 (I;III)*
1180	*lin-41(rrr3)/unc-29(e1072), lin-11(h1281) (I); rrrSi199 (II); rrrSi198 (IV)*
1534	*eri-1(mg366) (IV); rrrSi199 (II)*
1651	*prg-1(tm872) (I); rrrSi199 (II)*
1617	*mut-2(ne3370) (I); rrrSi199 (II)*
1616	*mut-7(ne4255) (III); rrrSi199 (II)*
1610	*drh-3(fj52) (I)/hT2 (I;III); rrrSi199 (II)*
1618	*avr-14(ad1302) drh-3(tm1217) (I)/hT2 (I;III) ;rrrSi199 (II)*
1596	*csr-1(tm892) (IV)/DnT1 (IV;V); rrrSi199 (II)*
WM191	MAGO12: *sago-2(tm894) ppw-1(tm914) ppw-2(tm110) F55A12.1(tm2686) R06C7.1(tm1414) I; Y49F6.1(tm1127) ZK1248.7(tm1113) F58G1.1(tm1019) II; C16C10.3(tm1200) K12B6.1(tm1195) III; T22H9.3(tm1186) V; R04A9.2(tm1116) X*

CGC or Ciosk lab number	**Genotype**
1646	*bqSi142[pBN20(unc-119(+) Pemr-1::emr-1::mCherry;)] (II)*
1634	*drh-3(rrr5) (I)/hT2 (I;III); bqSi142 (II); rrrSi198 (IV)*
1718	*drh-3(rrr5) (I)/hT2 (I;III); rrrSi318[Pvet-4::mCherry:h2b::tbb-2 3'UTR; unc-119(+)] (IV)*
1689	*csr-1(tm892) (IV); Pcsr-1::FLAG-CSR-1(DDH)::csr-1 3'UTR (II)*
1755	*csr-1(tm892) (IV)/DnT1 (IV;V); Pcsr-1::FLAG-CSR-1(AAA)::csr-1 3'UTR (II)*
1897	*drh-3(rrr2) (I)/hT2 (I;III); rrrSi321 [phsp-16.41::FLAG-GFP-LIN-41::lin-41 full-length 3'UTR, unc-119(+)] (IV)*
1777	*lin-41(rrr3) (I) /hT2 (I;III); GFP::CSR-1(CRISPR) (IV)*
1689	*csr-1(tm892) (IV); Pcsr-1::FLAG-CSR-1::csr-1 3'UTR (II)*
1591	*rrr6/mIn1 (II); rrrSi199 (II); rrrSi198 (IV); zuIs252; glo-1(zu391) (X)*
1184	*ran-2(rrr7) (III)/hT2 (I;III); rrrSi199 (II); rrrSi198 (IV); zuIs252; glo-1(zu391) (X)*
1528	*usp-48(rrr8) (I)/hT2 (I;III); rrrSi199 (II); rrrSi198 (IV); zuIs252; glo-1(zu391) (X)*

I thank Craig Mello for sharing several CSR-1 strains, including the CSR-1 slicer mutant, the endogenously GFP-tagged CSR-1 strain and the FLAG-tagged CSR-1 strain.

I thank Peter Askjaer for providing the EMR-1::mCherry strain.

Real-time quantitative PCR primer sequences

Primer	Sequence
act-1 FW	CTATGTTCCAGCCATCCTTCTTGG
act-1 RV	TGATCTTGATCTTCATGGTTGATGG
tbb-2 FW	GCTCATTCTCGGTTGTACCA
tbb-2 RV	TGGTGAGGGATACAAGATGG
vet-4 FW	AAGGATTTCACTGCTTGCTC
vet-4 RV	CGTCGTTTTCGATTTCTCCG
vet-6 FW	GTGCGAGACAAGAATGTAATCC
vet-6 RV	TTCTTGAACTCTTGGAACACAG
pes-10 FW	GCGATGATTTCATGATTTCCTG
pes-10 RV	AATTTCGTAGTCAATCTGCTCC
hlh-1 FW	ACGATTATGTGACTTCCTCTC
hlh-1 RV	GATGATCTCTATCGTCGTCC
unc-120 FW	GGGTATTATGAAGAAGGCATTCG
unc-120 RV	TGCATATGTGTAGACATGACCA
end-1 FW	GGGCAATACTTTGTTCAATCG
end-1 RV	GGATACTGTTGTGAGTAGCA
end-3 FW	GCCTATTAATGACCTCCAGC
end-3 RV	CCCGTCAATTGGTATCTCTG
pha-4 FW	CCAGAATTTCCTGAACAACAC
pha-4 RV	GTTGGTGGAGCTGTAAAGAG
elt-1 FW	ACAATTCTCAATTCAGCACG
elt-1 RV	GTTGCAGAGGTAGTTTCCTG

7 Acknowledgements

I want to thank

- my supervisor Rafal Ciosk for his guidance during my PhD, for giving me the freedom and trust that allowed me to explore things on my own, but for providing support when I asked for it

- my PhD committee members Susan Gasser and Christian Lehner for great input during our meetings, and Marc Bühler for joining the meetings and providing helpful suggestions, when my project suddenly turned to small RNAs

- everyone for making the FMI such a great place to do science: the facilities, especially the genomics facility and the worm facility, the media kitchen, the IT... Many thanks also to Elida, for being such a positive and supportive person in all organizational things during my PhD

- Michael and Dimos, who analyzed all the genome-wide data and always tried to make me as well understand what they do

- the Grosshans group for discussions and input during our joint lab meetings

- all current and previous members of the Ciosk lab: Anca, Andi, Balazs, Conny, Cris, Jane, Janosch, Jorge, Lukas, Pooja, Ricky, Sandra, Silvia, Tina, Vishal and Yanwu. A special thank goes to Cris for setting up the genetic screen and for teaching me a lot of things during my PhD. Many thanks also to Pooja for great teamwork on the CSR-1/LIN-41 project and helpful comments during writing the thesis. Many of you I see today more as friends than as colleagues!

- the lizard team Janine, Peter, Keith and Isabella, also for fun meetings outside of the FMI

- the PhD student representatives team for organizing great retreats, workshops, talks…and many more things together

- to my friends near and far, to Giovanna and Ivana for taking “PhD breaks” together, was it just a coffee break or a weekend in France!

- my family for their love and constant support

- Alex, for travelling countless times to Basel, for listening to me and believing in me…for everything!

8 References

Akkers RC, van Heeringen SJ, Jacobi UG, Janssen-Megens EM, Francoijs KJ, Stunnenberg HG, Veenstra GJ (2009) A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in Xenopus embryos. Developmental cell 17 (3):425-434. doi:10.1016/j.devcel.2009.08.005

Allen AK, Nesmith JE, Golden A (2014) An RNAi-Based Suppressor Screen Identifies Interactors of the Myt1 Ortholog of Caenorhabditis elegans. G3 4 (12):2329-2343. doi:10.1534/g3.114.013649

Aoki K, Moriguchi H, Yoshioka T, Okawa K, Tabara H (2007) In vitro analyses of the production and activity of secondary small interfering RNAs in C. elegans. The EMBO journal 26 (24):5007-5019. doi:10.1038/sj.emboj.7601910

Ashe A, Sapetschnig A, Weick EM, Mitchell J, Bagijn MP, Cording AC, Doebley AL, Goldstein LD, Lehrbach NJ, Le Pen J, Pintacuda G, Sakaguchi A, Sarkies P, Ahmed S, Miska EA (2012) piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell 150 (1):88-99. doi:10.1016/j.cell.2012.06.018

Askjaer P, Galy V, Hannak E, Mattaj IW (2002) Ran GTPase cycle and importins alpha and beta are essential for spindle formation and nuclear envelope assembly in living Caenorhabditis elegans embryos. Molecular biology of the cell 13 (12):4355-4370. doi:10.1091/mbc.E02-06-0346

Au KF, Jiang H, Lin L, Xing Y, Wong WH (2010) Detection of splice junctions from paired-end RNA-seq data by SpliceMap. Nucleic acids research 38 (14):4570-4578. doi:10.1093/nar/gkq211

Avgousti DC, Palani S, Sherman Y, Grishok A (2012) CSR-1 RNAi pathway positively regulates histone expression in C. elegans. The EMBO journal 31 (19):3821-3832. doi:10.1038/emboj.2012.216

Bagijn MP, Goldstein LD, Sapetschnig A, Weick EM, Bouasker S, Lehrbach NJ, Simard MJ, Miska EA (2012) Function, targets, and evolution of Caenorhabditis elegans piRNAs. Science 337 (6094):574-578. doi:10.1126/science.1220952

Barberan-Soler S, Fontrodona L, Ribo A, Lamm AT, Iannone C, Ceron J, Lehner B, Valcarcel J (2014) Co-option of the piRNA pathway for germline-specific alternative splicing of C. elegans TOR. Cell reports 8 (6):1609-1616. doi:10.1016/j.celrep.2014.08.016

Barckmann B, Simonelig M (2013) Control of maternal mRNA stability in germ cells and early embryos. Biochimica et biophysica acta 1829 (6-7):714-724. doi:10.1016/j.bbagrm.2012.12.011

Batista PJ, Ruby JG, Claycomb JM, Chiang R, Fahlgren N, Kasschau KD, Chaves DA, Gu W, Vasale JJ, Duan S, Conte D, Jr., Luo S, Schroth GP, Carrington JC, Bartel DP, Mello CC (2008) PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Molecular cell 31 (1):67-78. doi:10.1016/j.molcel.2008.06.002

Baugh LR, Hill AA, Claggett JM, Hill-Harfe K, Wen JC, Slonim DK, Brown EL, Hunter CP (2005) The homeodomain protein PAL-1 specifies a lineage-specific regulatory network in the C. elegans embryo. Development 132 (8):1843-1854. doi:10.1242/dev.01782

Baugh LR, Hill AA, Slonim DK, Brown EL, Hunter CP (2003) Composition and dynamics of the Caenorhabditis elegans early embryonic transcriptome. Development 130 (5):889-900

Bellier S, Chastant S, Adenot P, Vincent M, Renard JP, Bensaude O (1997) Nuclear translocation and carboxyl-terminal domain phosphorylation of RNA polymerase II delineate the two phases of zygotic gene activation in mammalian embryos. The EMBO journal 16 (20):6250-6262. doi:10.1093/emboj/16.20.6250

Benoit B, He CH, Zhang F, Votruba SM, Tadros W, Westwood JT, Smibert CA, Lipshitz HD, Theurkauf WE (2009) An essential role for the RNA-binding protein Smaug during the Drosophila maternal-to-zygotic transition. Development 136 (6):923-932. doi:10.1242/dev.031815

Biedermann B, Wright J, Senften M, Kalchhauser I, Sarathy G, Lee MH, Ciosk R (2009) Translational repression of cyclin E prevents precocious mitosis and embryonic gene activation during C. elegans meiosis. Developmental cell 17 (3):355-364. doi:10.1016/j.devcel.2009.08.003

Billi AC, Fischer SE, Kim JK (2014) Endogenous RNAi pathways in C. elegans. WormBook : the online review of C elegans biology:1-49. doi:10.1895/wormbook.1.170.1

Blackwell TK, Walker AK (2006) Transcription mechanisms. WormBook : the online review of C elegans biology:1-16. doi:10.1895/wormbook.1.121.1

Bogdanovic O, Long SW, van Heeringen SJ, Brinkman AB, Gomez-Skarmeta JL, Stunnenberg HG, Jones PL, Veenstra GJ (2011) Temporal uncoupling of the DNA methylome and transcriptional repression during embryogenesis. Genome research 21 (8):1313-1327. doi:10.1101/gr.114843.110

Brenner S (1974) The genetics of Caenorhabditis elegans. Genetics 77 (1):71-94

Brevini TA, Pennarossa G, Vanelli A, Maffei S, Gandolfi F (2012) Parthenogenesis in non-rodent species: developmental competence and differentiation plasticity. Theriogenology 77 (4):766-772. doi:10.1016/j.theriogenology.2011.11.010

Buckley BA, Burkhart KB, Gu SG, Spracklin G, Kershner A, Fritz H, Kimble J, Fire A, Kennedy S (2012) A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature 489 (7416):447-451. doi:10.1038/nature11352

Bultman SJ, Gebuhr TC, Pan H, Svoboda P, Schultz RM, Magnuson T (2006) Maternal BRG1 regulates zygotic genome activation in the mouse. Genes & development 20 (13):1744-1754. doi:10.1101/gad.1435106

Campbell AC, Updike DL (2015) CSR-1 and P granules suppress sperm-specific transcription in the C. elegans germline. Development 142 (10):1745-1755. doi:10.1242/dev.121434

Cantone I, Fisher AG (2013) Epigenetic programming and reprogramming during development. Nature structural & molecular biology 20 (3):282-289. doi:10.1038/nsmb.2489

Cecere G, Hoersch S, O'Keeffe S, Sachidanandam R, Grishok A (2014) Global effects of the CSR-1 RNA interference pathway on the transcriptional landscape. Nature structural & molecular biology 21 (4):358-365. doi:10.1038/nsmb.2801

Cecere G, Zheng GX, Mansisidor AR, Klymko KE, Grishok A (2012) Promoters recognized by forkhead proteins exist for individual 21U-RNAs. Molecular cell 47 (5):734-745. doi:10.1016/j.molcel.2012.06.021

Chang CC, Ma Y, Jacobs S, Tian XC, Yang X, Rasmussen TP (2005) A maternal store of macroH2A is removed from pronuclei prior to onset of somatic macroH2A expression in preimplantation embryos. Developmental biology 278 (2):367-380. doi:10.1016/j.ydbio.2004.11.032

Ciosk R, DePalma M, Priess JR (2006) Translational regulators maintain totipotency in the Caenorhabditis elegans germline. Science 311 (5762):851-853. doi:10.1126/science.1122491

Claycomb JM, Batista PJ, Pang KM, Gu W, Vasale JJ, van Wolfswinkel JC, Chaves DA, Shirayama M, Mitani S, Ketting RF, Conte D, Jr., Mello CC (2009) The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell 139 (1):123-134. doi:10.1016/j.cell.2009.09.014

Conine CC, Batista PJ, Gu W, Claycomb JM, Chaves DA, Shirayama M, Mello CC (2010) Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 107 (8):3588-3593. doi:10.1073/pnas.0911685107

Conine CC, Moresco JJ, Gu W, Shirayama M, Conte D, Jr., Yates JR, 3rd, Mello CC (2013) Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. Cell 155 (7):1532-1544. doi:10.1016/j.cell.2013.11.032

Cook MS, Munger SC, Nadeau JH, Capel B (2011) Regulation of male germ cell cycle arrest and differentiation by DND1 is modulated by genetic background. Development 138 (1):23-32. doi:10.1242/dev.057000

Crittenden SL, Bernstein DS, Bachorik JL, Thompson BE, Gallegos M, Petcherski AG, Moulder G, Barstead R, Wickens M, Kimble J (2002) A conserved RNA-binding protein controls germline stem cells in Caenorhabditis elegans. Nature 417 (6889):660-663. doi:10.1038/nature754

Cui M, Kim EB, Han M (2006) Diverse chromatin remodeling genes antagonize the Rb-involved SynMuv pathways in C. elegans. PLoS genetics 2 (5):e74. doi:10.1371/journal.pgen.0020074

De Renzis S, Elemento O, Tavazoie S, Wieschaus EF (2007) Unmasking activation of the zygotic genome using chromosomal deletions in the Drosophila embryo. PLoS biology 5 (5):e117. doi:10.1371/journal.pbio.0050117

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature genetics 43 (5):491-498. doi:10.1038/ng.806

Detwiler MR, Reuben M, Li X, Rogers E, Lin R (2001) Two zinc finger proteins, OMA-1 and OMA-2, are redundantly required for oocyte maturation in C. elegans. Developmental cell 1 (2):187-199

Dolci S, Campolo F, De Felici M (2015) Gonadal development and germ cell tumors in mouse and humans. Seminars in cell & developmental biology 45:114-123. doi:10.1016/j.semcdb.2015.10.002

Drake JW, Baltz RH (1976) The biochemistry of mutagenesis. Annual review of biochemistry 45:11-37. doi:10.1146/annurev.bi.45.070176.000303

Duchaine TF, Wohlschlegel JA, Kennedy S, Bei Y, Conte D, Jr., Pang K, Brownell DR, Harding S, Mitani S, Ruvkun G, Yates JR, 3rd, Mello CC (2006) Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell 124 (2):343-354. doi:10.1016/j.cell.2005.11.036

Dunican DS, Ruzov A, Hackett JA, Meehan RR (2008) xDnmt1 regulates transcriptional silencing in pre-MBT Xenopus embryos independently of its catalytic function. Development 135 (7):1295-1302. doi:10.1242/dev.016402

Edgar BA, Schubiger G (1986) Parameters controlling transcriptional activation during early Drosophila development. Cell 44 (6):871-877

Edgar LG, Wolf N, Wood WB (1994) Early transcription in Caenorhabditis elegans embryos. Development 120 (2):443-451

Ellis RE, Kimble J (1994) Control of germ cell differentiation in Caenorhabditis elegans. Ciba Foundation symposium 182:179-188; discussion 189-192

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391 (6669):806-811. doi:10.1038/35888

Fischer SE, Montgomery TA, Zhang C, Fahlgren N, Breen PC, Hwang A, Sullivan CM, Carrington JC, Ruvkun G (2011) The ERI-6/7 helicase acts at the first stage of an siRNA amplification pathway that targets recent gene duplications. PLoS genetics 7 (11):e1002369. doi:10.1371/journal.pgen.1002369

Flemr M, Malik R, Franke V, Nejepinska J, Sedlacek R, Vlahovicek K, Svoboda P (2013) A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. Cell 155 (4):807-816. doi:10.1016/j.cell.2013.10.001

Francis R, Barton MK, Kimble J, Schedl T (1995) gld-1, a tumor suppressor gene required for oocyte development in Caenorhabditis elegans. Genetics 139 (2):579-606

Friend K, Campbell ZT, Cooke A, Kroll-Conner P, Wickens MP, Kimble J (2012) A conserved PUF-Ago-eEF1A complex attenuates translation elongation. Nature structural & molecular biology 19 (2):176-183. doi:10.1038/nsmb.2214

Gaidatzis D, Lerch A, Hahne F, Stadler MB (2015) QuasR: quantification and annotation of short reads in R. Bioinformatics 31 (7):1130-1132. doi:10.1093/bioinformatics/btu781

Gerson-Gurwitz A, Wang S, Sathe S, Green R, Yeo GW, Oegema K, Desai A (2016) A Small RNA-Catalytic Argonaute Pathway Tunes Germline Transcript Levels to Ensure Embryonic Divisions. Cell 165 (2):396-409. doi:10.1016/j.cell.2016.02.040

Gu W, Claycomb JM, Batista PJ, Mello CC, Conte D (2011) Cloning Argonaute-associated small RNAs from Caenorhabditis elegans. Methods in molecular biology 725:251-280. doi:10.1007/978-1-61779-046-1_17

Gu W, Shirayama M, Conte D, Jr., Vasale J, Batista PJ, Claycomb JM, Moresco JJ, Youngman EM, Keys J, Stoltz MJ, Chen CC, Chaves DA, Duan S, Kasschau KD, Fahlgren N, Yates JR, 3rd, Mitani S, Carrington JC, Mello CC (2009) Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. Molecular cell 36 (2):231-244. doi:10.1016/j.molcel.2009.09.020

Guang S, Bochner AF, Burkhart KB, Burton N, Pavelec DM, Kennedy S (2010) Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. Nature 465 (7301):1097-1101. doi:10.1038/nature09095

Gurdon JB, Uehlinger V (1966) "Fertile" intestine nuclei. Nature 210 (5042):1240-1241

Guven-Ozkan T, Nishi Y, Robertson SM, Lin R (2008) Global transcriptional repression in C. elegans germline precursors by regulated sequestration of TAF-4. Cell 135 (1):149-160. doi:10.1016/j.cell.2008.07.040

Haberle V, Li N, Hadzhiev Y, Plessy C, Previti C, Nepal C, Gehrig J, Dong X, Akalin A, Suzuki AM, van IWF, Armant O, Ferg M, Strahle U, Carninci P, Muller F, Lenhard B (2014) Two independent transcription initiation codes overlap on vertebrate core promoters. Nature 507 (7492):381-385. doi:10.1038/nature12974

Han T, Manoharan AP, Harkins TT, Bouffard P, Fitzpatrick C, Chu DS, Thierry-Mieg D, Thierry-Mieg J, Kim JK (2009) 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 106 (44):18674-18679. doi:10.1073/pnas.0906378106

Harrison MM, Li XY, Kaplan T, Botchan MR, Eisen MB (2011) Zelda binding in the early Drosophila melanogaster embryo marks regions subsequently activated at the maternal-to-zygotic transition. PLoS genetics 7 (10):e1002266. doi:10.1371/journal.pgen.1002266

Hendriks GJ, Gaidatzis D, Aeschimann F, Grosshans H (2014) Extensive oscillatory gene expression during C. elegans larval development. Molecular cell 53 (3):380-392. doi:10.1016/j.molcel.2013.12.013

Heyn P, Kircher M, Dahl A, Kelso J, Tomancak P, Kalinka AT, Neugebauer KM (2014) The earliest transcribed zygotic genes are short, newly evolved, and different across species. Cell reports 6 (2):285-292. doi:10.1016/j.celrep.2013.12.030

Hunter CP, Kenyon C (1996) Spatial and temporal controls target pal-1 blastomere-specification activity to a single blastomere lineage in C. elegans embryos. Cell 87 (2):217-226

Hussain SA, Ma YT, Palmer DH, Hutton P, Cullen MH (2008) Biology of testicular germ cell tumors. Expert review of anticancer therapy 8 (10):1659-1673. doi:10.1586/14737140.8.10.1659

Keall R, Whitelaw S, Pettitt J, Muller B (2007) Histone gene expression and histone mRNA 3' end structure in Caenorhabditis elegans. BMC molecular biology 8:51. doi:10.1186/1471-2199-8-51

Kedde M, Strasser MJ, Boldajipour B, Oude Vrielink JA, Slanchev K, le Sage C, Nagel R, Voorhoeve PM, van Duijse J, Orom UA, Lund AH, Perrakis A, Raz E, Agami R (2007) RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. Cell 131 (7):1273-1286. doi:10.1016/j.cell.2007.11.034

Kelly WG, Schaner CE, Dernburg AF, Lee MH, Kim SK, Villeneuve AM, Reinke V (2002) X-chromosome silencing in the germline of C. elegans. Development 129 (2):479-492

Kennedy S, Wang D, Ruvkun G (2004) A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans. Nature 427 (6975):645-649. doi:10.1038/nature02302

Kerr CL, Shamblott MJ, Gearhart JD (2006) Pluripotent stem cells from germ cells. Methods in enzymology 419:400-426. doi:10.1016/S0076-6879(06)19016-3

Kim S, Spike C, Greenstein D (2013) Control of oocyte growth and meiotic maturation in Caenorhabditis elegans. Advances in experimental medicine and biology 757:277-320. doi:10.1007/978-1-4614-4015-4_10

Kimble JE, White JG (1981) On the control of germ cell development in Caenorhabditis elegans. Developmental biology 81 (2):208-219

Kimelman D, Kirschner M, Scherson T (1987) The events of the midblastula transition in Xenopus are regulated by changes in the cell cycle. Cell 48 (3):399-407

Kono T, Obata Y, Wu Q, Niwa K, Ono Y, Yamamoto Y, Park ES, Seo JS, Ogawa H (2004) Birth of parthenogenetic mice that can develop to adulthood. Nature 428 (6985):860-864. doi:10.1038/nature02402

Kumagai A, Dunphy WG (1991) The cdc25 protein controls tyrosine dephosphorylation of the cdc2 protein in a cell-free system. Cell 64 (5):903-914

LaMarca MJ, Fidler MC, Smith LD, Keem K (1975) Hormonal effects on RNA synthesis by stage 6 oocytes of Xenopus laevis. Developmental biology 47 (2):384-393

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10 (3):R25. doi:10.1186/gb-2009-10-3-r25

Lee HC, Gu W, Shirayama M, Youngman E, Conte D, Jr., Mello CC (2012) C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. Cell 150 (1):78-87. doi:10.1016/j.cell.2012.06.016

Lee MT, Bonneau AR, Giraldez AJ (2014) Zygotic genome activation during the maternal-to-zygotic transition. Annual review of cell and developmental biology 30:581-613. doi:10.1146/annurev-cellbio-100913-013027

Lee MT, Bonneau AR, Takacs CM, Bazzini AA, DiVito KR, Fleming ES, Giraldez AJ (2013) Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. Nature 503 (7476):360-364. doi:10.1038/nature12632

Lee RC, Hammell CM, Ambros V (2006) Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. Rna 12 (4):589-597. doi:10.1261/rna.2231506

Leichsenring M, Maes J, Mossner R, Driever W, Onichtchouk D (2013) Pou5f1 transcription factor controls zygotic gene activation in vertebrates. Science 341 (6149):1005-1009. doi:10.1126/science.1242527

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25 (14):1754-1760. doi:10.1093/bioinformatics/btp324

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25 (16):2078-2079. doi:10.1093/bioinformatics/btp352

Li XY, Harrison MM, Villalta JE, Kaplan T, Eisen MB (2014) Establishment of regions of genomic activity during the Drosophila maternal to zygotic transition. eLife 3. doi:10.7554/eLife.03737

Liang HL, Nien CY, Liu HY, Metzstein MM, Kirov N, Rushlow C (2008) The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. Nature 456 (7220):400-403. doi:10.1038/nature07388

Lin R, Hill RJ, Priess JR (1998) POP-1 and anterior-posterior fate decisions in C. elegans embryos. Cell 92 (2):229-239

Lindeman LC, Andersen IS, Reiner AH, Li N, Aanes H, Ostrup O, Winata C, Mathavan S, Muller F, Alestrom P, Collas P (2011) Prepatterning of developmental gene expression by modified histones before zygotic genome activation. Developmental cell 21 (6):993-1004. doi:10.1016/j.devcel.2011.10.008

Loedige I, Stotz M, Qamar S, Kramer K, Hennig J, Schubert T, Loffler P, Langst G, Merkl R, Urlaub H, Meister G (2014) The NHL domain of BRAT is an RNA-binding domain that directly contacts the hunchback mRNA for regulation. Genes & development 28 (7):749-764. doi:10.1101/gad.236513.113

Luteijn MJ, van Bergeijk P, Kaaij LJ, Almeida MV, Roovers EF, Berezikov E, Ketting RF (2012) Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. The EMBO journal 31 (16):3422-3430. doi:10.1038/emboj.2012.213

Ma J, Flemr M, Stein P, Berninger P, Malik R, Zavolan M, Svoboda P, Schultz RM (2010) MicroRNA activity is suppressed in mouse oocytes. Current biology : CB 20 (3):265-270. doi:10.1016/j.cub.2009.12.042

Maine EM, Hauth J, Ratliff T, Vought VE, She X, Kelly WG (2005) EGO-1, a putative RNA-dependent RNA polymerase, is required for heterochromatin assembly on unpaired dna during C. elegans meiosis. Current biology : CB 15 (21):1972-1978. doi:10.1016/j.cub.2005.09.049

Masui Y (2001) From oocyte maturation to the in vitro cell cycle: the history of discoveries of Maturation-Promoting Factor (MPF) and Cytostatic Factor (CSF). Differentiation; research in biological diversity 69 (1):1-17. doi:10.1046/j.1432-0436.2001.690101.x

McCarter J, Bartlett B, Dang T, Schedl T (1999) On the control of oocyte meiotic maturation and ovulation in Caenorhabditis elegans. Developmental biology 205 (1):111-128. doi:10.1006/dbio.1998.9109

McKim KS, Peters K, Rose AM (1993) Two types of sites required for meiotic chromosome pairing in Caenorhabditis elegans. Genetics 134 (3):749-768

Mello CC, Draper BW, Krause M, Weintraub H, Priess JR (1992) The pie-1 and mex-1 genes and maternal control of blastomere identity in early C. elegans embryos. Cell 70 (1):163-176

Merritt C, Rasoloson D, Ko D, Seydoux G (2008) 3' UTRs are the primary regulators of gene expression in the C. elegans germline. Current biology : CB 18 (19):1476-1482. doi:10.1016/j.cub.2008.08.013

Mikeladze-Dvali T, von Tobel L, Strnad P, Knott G, Leonhardt H, Schermelleh L, Gonczy P (2012) Analysis of centriole elimination during C. elegans oogenesis. Development 139 (9):1670-1679. doi:10.1242/dev.075440

Miller MA, Nguyen VQ, Lee MH, Kosinski M, Schedl T, Caprioli RM, Greenstein D (2001) A sperm cytoskeletal protein that signals oocyte meiotic maturation and ovulation. Science 291 (5511):2144-2147. doi:10.1126/science.1057586

Moore GP, Lintern-Moore S, Peters H, Faber M (1974) RNA synthesis in the mouse oocyte. The Journal of cell biology 60 (2):416-422

Morales-Martinez A, Dobrzynska A, 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):1090-1096. doi:10.1242/jcs.164202

Mueller PR, Coleman TR, Dunphy WG (1995) Cell cycle regulation of a Xenopus Wee1-like kinase. Molecular biology of the cell 6 (1):119-134

Murchison EP, Stein P, Xuan Z, Pan H, Zhang MQ, Schultz RM, Hannon GJ (2007) Critical roles for Dicer in the female germline. Genes & development 21 (6):682-693. doi:10.1101/gad.1521307

Nepal C, Hadzhiev Y, Previti C, Haberle V, Li N, Takahashi H, Suzuki AM, Sheng Y, Abdelhamid RF, Anand S, Gehrig J, Akalin A, Kockx CE, van der Sloot AA, van Ijcken WF, Armant O, Rastegar S, Watson C, Strahle U, Stupka E, Carninci P, Lenhard B, Muller F (2013) Dynamic regulation of the transcription initiation landscape at single nucleotide resolution during vertebrate embryogenesis. Genome research 23 (11):1938-1950. doi:10.1101/gr.153692.112

Newport J, Kirschner M (1982) A major developmental transition in early Xenopus embryos: I. characterization and timing of cellular changes at the midblastula stage. Cell 30 (3):675-686

Nien CY, Liang HL, Butcher S, Sun Y, Fu S, Gocha T, Kirov N, Manak JR, Rushlow C (2011) Temporal coordination of gene networks by Zelda in the early Drosophila embryo. PLoS genetics 7 (10):e1002339. doi:10.1371/journal.pgen.1002339

Papaevgeniou N, Chondrogianni N (2014) The ubiquitin proteasome system in Caenorhabditis elegans and its regulation. Redox biology 2:333-347. doi:10.1016/j.redox.2014.01.007

Paranjpe SS, Veenstra GJ (2015) Establishing pluripotency in early development. Biochimica et biophysica acta 1849 (6):626-636. doi:10.1016/j.bbagrm.2015.03.006

Pasque V, Jullien J, Miyamoto K, Halley-Stott RP, Gurdon JB (2011) Epigenetic factors influencing resistance to nuclear reprogramming. Trends in genetics : TIG 27 (12):516-525. doi:10.1016/j.tig.2011.08.002

Patel T, Tursun B, Rahe DP, Hobert O (2012) Removal of Polycomb Repressive Complex 2 Makes C. elegans Germ Cells Susceptible to Direct Conversion into Specific Somatic Cell Types. Cell reports 2 (5):1178-1186. doi:10.1016/j.celrep.2012.09.020

Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D, Knowles BB (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. Developmental cell 7 (4):597-606. doi:10.1016/j.devcel.2004.09.004

Perez-Montero S, Carbonell A, Moran T, Vaquero A, Azorin F (2013) The embryonic linker histone H1 variant of Drosophila, dBigH1, regulates zygotic genome activation. Developmental cell 26 (6):578-590. doi:10.1016/j.devcel.2013.08.011

Phillips CM, Montgomery TA, Breen PC, Ruvkun G (2012) MUT-16 promotes formation of perinuclear mutator foci required for RNA silencing in the C. elegans germline. Genes & development 26 (13):1433-1444. doi:10.1101/gad.193904.112

Qiao L, Lissemore JL, Shu P, Smardon A, Gelber MB, Maine EM (1995) Enhancers of glp-1, a gene required for cell-signaling in Caenorhabditis elegans, define a set of genes required for germline development. Genetics 141 (2):551-569

Robert VJ, Mercier MG, Bedet C, Janczarski S, Merlet J, Garvis S, Ciosk R, Palladino F (2014) The SET-2/SET1 histone H3K4 methyltransferase maintains pluripotency in the Caenorhabditis elegans germline. Cell reports 9 (2):443-450. doi:10.1016/j.celrep.2014.09.018

Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell 127 (6):1193-1207. doi:10.1016/j.cell.2006.10.040

Scheckel C, Gaidatzis D, Wright JE, Ciosk R (2012) Genome-wide analysis of GLD-1-mediated mRNA regulation suggests a role in mRNA storage. PLoS genetics 8 (5):e1002742. doi:10.1371/journal.pgen.1002742

Schisa JA, Pitt JN, Priess JR (2001) Analysis of RNA associated with P granules in germ cells of C. elegans adults. Development 128 (8):1287-1298

Schulz KN, Bondra ER, Moshe A, Villalta JE, Lieb JD, Kaplan T, McKay DJ, Harrison MM (2015) Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early Drosophila embryo. Genome research 25 (11):1715-1726. doi:10.1101/gr.192682.115

Sega GA (1984) A review of the genetic effects of ethyl methanesulfonate. Mutation research 134 (2-3):113-142

Seth M, Shirayama M, Gu W, Ishidate T, Conte D, Jr., Mello CC (2013) The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. Developmental cell 27 (6):656-663. doi:10.1016/j.devcel.2013.11.014

Seydoux G, Dunn MA (1997) Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of Caenorhabditis elegans and Drosophila melanogaster. Development 124 (11):2191-2201

Seydoux G, Fire A (1994) Soma-germline asymmetry in the distributions of embryonic RNAs in Caenorhabditis elegans. Development 120 (10):2823-2834

She X, Xu X, Fedotov A, Kelly WG, Maine EM (2009) Regulation of heterochromatin assembly on unpaired chromosomes during Caenorhabditis elegans meiosis by components of a small RNA-mediated pathway. PLoS genetics 5 (8):e1000624. doi:10.1371/journal.pgen.1000624

Shermoen AW, O'Farrell PH (1991) Progression of the cell cycle through mitosis leads to abortion of nascent transcripts. Cell 67 (2):303-310

Shimada M, Kawahara H, Doi H (2002) Novel family of CCCH-type zinc-finger proteins, MOE-1, -2 and -3, participates in C. elegans oocyte maturation. Genes to cells : devoted to molecular & cellular mechanisms 7 (9):933-947

Shirayama M, Seth M, Lee HC, Gu W, Ishidate T, Conte D, Jr., Mello CC (2012) piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell 150 (1):65-77. doi:10.1016/j.cell.2012.06.015

Smardon A, Spoerke JM, Stacey SC, Klein ME, Mackin N, Maine EM (2000) EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Current biology : CB 10 (4):169-178

Smith AG (2001) Embryo-derived stem cells: of mice and men. Annual review of cell and developmental biology 17:435-462. doi:10.1146/annurev.cellbio.17.1.435

Spike CA, Coetzee D, Eichten C, Wang X, Hansen D, Greenstein D (2014) The TRIM-NHL protein LIN-41 and the OMA RNA-binding proteins antagonistically control the prophase-to-metaphase transition and growth of Caenorhabditis elegans oocytes. Genetics 198 (4):1535-1558. doi:10.1534/genetics.114.168831

Stancheva I, Meehan RR (2000) Transient depletion of xDnmt1 leads to premature gene activation in Xenopus embryos. Genes & development 14 (3):313-327

Stein P, Rozhkov NV, Li F, Cardenas FL, Davydenko O, Vandivier LE, Gregory BD, Hannon GJ, Schultz RM (2015) Essential Role for endogenous siRNAs during meiosis in mouse oocytes. PLoS genetics 11 (2):e1005013. doi:10.1371/journal.pgen.1005013

Strome S, Updike D (2015) Specifying and protecting germ cell fate. Nature reviews Molecular cell biology 16 (7):406-416. doi:10.1038/nrm4009

Suh N, Baehner L, Moltzahn F, Melton C, Shenoy A, Chen J, Blelloch R (2010) MicroRNA function is globally suppressed in mouse oocytes and early embryos. Current biology : CB 20 (3):271-277. doi:10.1016/j.cub.2009.12.044

Suh N, Blelloch R (2011) Small RNAs in early mammalian development: from gametes to gastrulation. Development 138 (9):1653-1661. doi:10.1242/dev.056234

Swarts DC, Makarova K, Wang Y, Nakanishi K, Ketting RF, Koonin EV, Patel DJ, van der Oost J (2014) The evolutionary journey of Argonaute proteins. Nature structural & molecular biology 21 (9):743-753. doi:10.1038/nsmb.2879

Tadros W, Goldman AL, Babak T, Menzies F, Vardy L, Orr-Weaver T, Hughes TR, Westwood JT, Smibert CA, Lipshitz HD (2007) SMAUG is a major regulator of maternal mRNA destabilization in Drosophila and its translation is activated by the PAN GU kinase. Developmental cell 12 (1):143-155. doi:10.1016/j.devcel.2006.10.005

Tadros W, Lipshitz HD (2009) The maternal-to-zygotic transition: a play in two acts. Development 136 (18):3033-3042. doi:10.1242/dev.033183

Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126 (4):663-676. doi:10.1016/j.cell.2006.07.024

Tang F, Kaneda M, O'Carroll D, Hajkova P, Barton SC, Sun YA, Lee C, Tarakhovsky A, Lao K, Surani MA (2007) Maternal microRNAs are essential for mouse zygotic development. Genes & development 21 (6):644-648. doi:10.1101/gad.418707

ten Bosch JR, Benavides JA, Cline TW (2006) The TAGteam DNA motif controls the timing of Drosophila pre-blastoderm transcription. Development 133 (10):1967-1977. doi:10.1242/dev.02373

Thivierge C, Makil N, Flamand M, Vasale JJ, Mello CC, Wohlschlegel J, Conte D, Jr., Duchaine TF (2011) Tudor domain ERI-5 tethers an RNA-dependent RNA polymerase to DCR-1 to potentiate endo-RNAi. Nature structural & molecular biology 19 (1):90-97. doi:10.1038/nsmb.2186

Tocchini C (2015) Control of pluripotency during the oocyte-to-embryo transition in *Caenorhabditis elegans* PhD thesis, University of Basel, Switzerland

Tocchini C, Keusch JJ, Miller SB, Finger S, Gut H, Stadler MB, Ciosk R (2014) The TRIM-NHL protein LIN-41 controls the onset of developmental plasticity in Caenorhabditis elegans. PLoS genetics 10 (8):e1004533. doi:10.1371/journal.pgen.1004533

Tursun B, Patel T, Kratsios P, Hobert O (2011) Direct conversion of C. elegans germ cells into specific neuron types. Science 331 (6015):304-308. doi:10.1126/science.1199082

Ulbright TM (2005) Germ cell tumors of the gonads: a selective review emphasizing problems in differential diagnosis, newly appreciated, and controversial issues. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc 18 Suppl 2:S61-79. doi:10.1038/modpathol.3800310

Updike DL, Knutson AK, Egelhofer TA, Campbell AC, Strome S (2014) Germ-granule components prevent somatic development in the C. elegans germline. Current biology : CB 24 (9):970-975. doi:10.1016/j.cub.2014.03.015

Updike DL, Strome S (2009) A genomewide RNAi screen for genes that affect the stability, distribution and function of P granules in Caenorhabditis elegans. Genetics 183 (4):1397-1419. doi:10.1534/genetics.109.110171

van Wolfswinkel JC, Claycomb JM, Batista PJ, Mello CC, Berezikov E, Ketting RF (2009) CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell 139 (1):135-148. doi:10.1016/j.cell.2009.09.012

Vasale JJ, Gu W, Thivierge C, Batista PJ, Claycomb JM, Youngman EM, Duchaine TF, Mello CC, Conte D, Jr. (2010) Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. Proceedings of the National Academy of Sciences of the United States of America 107 (8):3582-3587. doi:10.1073/pnas.0911908107

Vastenhouw NL, Zhang Y, Woods IG, Imam F, Regev A, Liu XS, Rinn J, Schier AF (2010) Chromatin signature of embryonic pluripotency is established during genome activation. Nature 464 (7290):922-926. doi:10.1038/nature08866

Von Stetina JR, Orr-Weaver TL (2011) Developmental control of oocyte maturation and egg activation in metazoan models. Cold Spring Harbor perspectives in biology 3 (10):a005553. doi:10.1101/cshperspect.a005553

Voronina E, Paix A, Seydoux G (2012) The P granule component PGL-1 promotes the localization and silencing activity of the PUF protein FBF-2 in germline stem cells. Development 139 (20):3732-3740. doi:10.1242/dev.083980

Vought VE, Ohmachi M, Lee MH, Maine EM (2005) EGO-1, a putative RNA-directed RNA polymerase, promotes germline proliferation in parallel with GLP-1/notch signaling and regulates the spatial organization of nuclear pore complexes and germline P granules in Caenorhabditis elegans. Genetics 170 (3):1121-1132. doi:10.1534/genetics.105.042135

Walker AK, Boag PR, Blackwell TK (2007) Transcription reactivation steps stimulated by oocyte maturation in C. elegans. Developmental biology 304 (1):382-393. doi:10.1016/j.ydbio.2006.12.039

Walser CB, Lipshitz HD (2011) Transcript clearance during the maternal-to-zygotic transition. Curr Opin Genet Dev 21 (4):431-443. doi:10.1016/j.gde.2011.03.003

Wang G, Reinke V (2008) A C. elegans Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. Current biology : CB 18 (12):861-867. doi:10.1016/j.cub.2008.05.009

Wedeles CJ, Wu MZ, Claycomb JM (2013a) A multitasking Argonaute: exploring the many facets of C. elegans CSR-1. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology 21 (6-7):573-586. doi:10.1007/s10577-013-9383-7

Wedeles CJ, Wu MZ, Claycomb JM (2013b) Protection of germline gene expression by the C. elegans Argonaute CSR-1. Developmental cell 27 (6):664-671. doi:10.1016/j.devcel.2013.11.016

Weismann A (1893) The Germ-Plasm: A Theory of Heredity. Charles Scribner's Sons

Wulczyn FG, Cuevas E, Franzoni E, Rybak A (2011) miRNAs Need a Trim : Regulation of miRNA Activity by Trim-NHL Proteins. Advances in experimental medicine and biology 700:85-105. doi:10.1007/978-1-4419-7823-3_9

Yigit E, Batista PJ, Bei Y, Pang KM, Chen CC, Tolia NH, Joshua-Tor L, Mitani S, Simard MJ, Mello CC (2006) Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell 127 (4):747-757. doi:10.1016/j.cell.2006.09.033

Youngren KK, Coveney D, Peng X, Bhattacharya C, Schmidt LS, Nickerson ML, Lamb BT, Deng JM, Behringer RR, Capel B, Rubin EM, Nadeau JH, Matin A (2005) The Ter mutation in the dead end gene causes germ cell loss and testicular germ cell tumours. Nature 435 (7040):360-364. doi:10.1038/nature03595

Zhang C, Montgomery TA, Gabel HW, Fischer SE, Phillips CM, Fahlgren N, Sullivan CM, Carrington JC, Ruvkun G (2011) mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 108 (4):1201-1208. doi:10.1073/pnas.1018695108

Zhang F, Barboric M, Blackwell TK, Peterlin BM (2003) A model of repression: CTD analogs and PIE-1 inhibit transcriptional elongation by P-TEFb. Genes & development 17 (6):748-758. doi:10.1101/gad.1068203

Zong J, Yao X, Yin J, Zhang D, Ma H (2009) Evolution of the RNA-dependent RNA polymerase (RdRP) genes: duplications and possible losses before and after the divergence of major eukaryotic groups. Gene 447 (1):29-39. doi:10.1016/j.gene.2009.07.004

Zurita M, Reynaud E, Aguilar-Fuentes J (2008) From the beginning: the basal transcription machinery and onset of transcription in the early animal embryo. Cellular and molecular life sciences : CMLS 65 (2):212-227. doi:10.1007/s00018-007-7295-4

Zuryn S, Le Gras S, Jamet K, Jarriault S (2010) A strategy for direct mapping and identification of mutations by whole-genome sequencing. Genetics 186 (1):427-430. doi:10.1534/genetics.110.119230