The mRNA silencing dual activity of the RNA-binding protein LIN41 and the role of its targets in *C. elegans* development

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Florian Aeschimann
aus Trachselwald (BE)

Basel, 2017

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Prof. Dr. Mihaela Zavolan, Dr. Helge Großhans and Prof. Dr. Oliver Mühlemann.

Basel, den 24.5.2016

Prof. Dr. Jörg Schibler
(Dekan)

Summary

The microRNA *let-7* is an ancient and fundamental regulator of stem cell self-renewal and differentiation in animals. In *Caenorhabditis elegans* (*C. elegans*), where *let-7* was discovered, it controls self-renewal of stem-cell like cells in the worm skin, the seam cells. In this work, we find that *let-7* controls seam cell self-renewal exclusively through one key target, the RNA-binding protein LIN41. Interestingly, LIN41 is a conserved *let-7* target, and was shown to regulate developmental switches from a self-renewal to a differentiation program in mouse and human embryonic stem cells, as well as in mouse neural progenitor cells. Despite its prominent role in stem cell development, the physiologically relevant mRNA targets of LIN41, to which it binds to regulate stem cell development, have been unknown in both mammals and *C. elegans*. LIN41 is reported to repress target mRNAs by inducing mRNA degradation, but has also been speculated to be a translational repressor.

In this work, we establish and optimize the ribosome profiling technique for *C. elegans*, to subsequently study gene expression changes at the level of mRNA abundance and translation, in mutant worms with abnormally high LIN41 levels. With these experiments, performed *in vivo* in a whole organism, we identify candidate physiological LIN41 target mRNAs. Of those, we confirm four as direct LIN41 targets, as they are enriched in RNA co-immunoprecipitations of LIN41. Surprisingly, the repression mechanisms differ among the four targets, with LIN41 destabilizing *mab-10*, *mab-3*, and *dmd-3* transcripts, but translationally repressing the *lin-29A* mRNA. Consistent with its few reported mammalian targets, LIN41 silences three targets by binding to their 3'UTRs, but unexpectedly, it represses *lin-29A* through its 5'UTR, thereby achieving isoform-specific regulation. Strikingly, we find that it is the location of the binding site that instructs the silencing mechanism, with LIN41 bound to the 5'UTR causing translational repression, while causing mRNA degradation when bound to the 3'UTR. Furthermore, we identify the transcription factor *lin-29A* and its co-factor *mab-10* as the key targets of LIN41 for controlling self-renewal of seam cells. Because their mammalian homologs, EGR and NAB proteins, are crucial regulators of stem cell proliferation and differentiation, we hypothesize that the uncovered pathway to control seam cell self-renewal is conserved in mammalian stem cells, potentially including a LIN41-mediated regulation of *EGR* and *NAB* mRNAs.

While further characterizing the four LIN41 targets, we discover that they also explain the defects of *lin-41* mutants in sexual organ development. Thus, LIN41 controls the morphogenesis of the vulva through *lin-29A* and *mab-10*, and the maturation of the male tail tip through *mab-3* and *dmd-3*. Moreover, we find that the developmental transition from a larval to an adult *C. elegans* epidermis depends on the upregulation of both isoforms of LIN-29. To inhibit a premature expression, *lin-29A* is regulated by LIN41, while the expression of *lin-29B* is controlled through the transcription factor HBL-1.

Table of contents

1 Introduction

1.1 Post-transcriptional gene regulation by RNA-binding proteins and microRNAs

When organisms progress through development or respond to environmental stimuli, their cells change the expression of gene products. To achieve an accurate regulation of gene expression in both time and space, cells use different mechanisms to increase or decrease the levels of specific proteins or RNAs. Generally, these mechanisms either modulate the rate of transcription, i.e. the production of RNA from a DNA template, or of a post-transcriptional process, i.e. one of the various steps in the life of a gene product thereafter.

Protein production from messenger RNAs (mRNAs) in cells is heavily controlled on a post-transcriptional level (Schwanhäusser et al., 2011, 2013). This control can occur both on a global level, for example by regulating the global activity of translation, and on the level of individual mRNAs, through the regulation of transcript abundance or translational activity. Post-transcriptional regulation of specific mRNAs is commonly achieved by sequence-specific binding of either microRNAs (miRNAs) or RNA-binding proteins (RBPs), with RBPs sometimes recognizing a specific RNA structure instead of a linear sequence motif. Most RBPs contain an RNA-binding protein domain that specifically recognizes an RNA structure or sequence motif. One exception is the family of Argonaute (AGO) proteins that need to assemble with a miRNA to be able to bind to target mRNAs. Vice versa, post-transcriptional regulation by a miRNA is dependent on its assembly with an AGO protein. Therefore, the miRNA-AGO complex can be looked at as an RBP with the miRNA as its RNA-binding domain, and, as described below, it employs similar principles as other RBPs for post-transcriptional gene regulation.

RBPs can post-transcriptionally modulate gene expression at many different levels in the life of an mRNA, for example by regulation of alternative splicing, polyadenylation, RNA editing, nuclear export, localization, translation or turnover. Eventually, in order to control the amount of protein production from individual mRNAs in the cytoplasm, RBPs normally regulate the rates of translation and mRNA degradation. Thereby, some RBPs like AGO proteins influence both translation and degradation, while other RPBs can elicit changes in translational activity without changes in mRNA turnover and vice versa (although indirectly, degradation of the mRNA template always eventually inhibits translation). For the vast majority of RBPs regulating gene expression at the post-transcriptional level, their exact mechanisms of action are unclear, with much less mechanistic details known compared to, for example, AGO proteins.

1.1.1 Regulation of mRNA abundance through RBPs

Steady state levels of mRNAs are defined by their rates of synthesis and decay. Thus, to modify the abundance of target mRNAs at the post-transcriptional level, RBPs have to change the rate of their degradation. The decay rates have been shown to differ up to 100-fold amongst mRNAs, a range that can be exploited by RBPs through stabilization or de-stabilization of mRNAs (Perez-Ortin et al., 2013). To do so, RBPs can either protect mRNAs from, or expose mRNAs to, the cellular RNA degradation machineries. The major RNA degradation machineries in eukaryotic cells act through exonuclease activities (Siwaszek et al., 2014), and mRNAs are protected from those by a special structures on each end, a 5' m^7GpppN cap structure and a 3' poly(A) tail. Both 5' cap and poly(A) tail are coated with proteins specifically recognizing these structures, named cap-binding proteins (CBPs) and poly(A)-binding proteins (PABPs), respectively. The decay of mRNAs commonly starts by removing one of the two protective structures, and RBPs can act by accelerating or slowing down this removal. The most conserved and best studied proteins to remove poly(A) tails are the deadenylase complexes CCR4-NOT and PAN2-PAN3, as well as the poly(A)-specific ribonuclease PARN, while decapping is usually performed by the DCP1/DCP2 heterodimer. Exposed mRNA ends are mainly degraded by the 5'-to-3' exonuclease XRN1 or the cytoplasmic exosome protein complex harboring 3'-to-5' exonuclease activity. For efficient degradation of target mRNAs, RBPs can directly interact with some of the above-mentioned factors of mRNA decay pathways or, in the extreme case, even have RNase activity themselves.

1.1.2 Regulation of translation through RBPs

In order to modify the translation rate of target mRNAs, RBPs can in principle modulate any step of the translation process, which can be conceptually divided into three consecutive phases: Initiation, elongation and termination. Canonical translation initiation (Jackson et al., 2010) is dependent on the 5' m^7GpppN cap structure. In the cytoplasm, the cap of most translated mRNAs is bound to the cap-binding complex eIF4F, consisting of the three proteins eIF4E, eIF4G and eIF4A. Within this complex, eIF4E directly binds to the cap structure, eIF4G bridges the eIF4E-eIF4A interaction and acts as a scaffold to recruit other initiation factors and eIF4A is an RNA helicase thought to unwind secondary structures of the RNA in order to allow the ribosomes to land. Another initiation factor, eIF3, associates with the small ribosomal subunit and other initiation factors to form the 43S-preinitiation complex. Because eIF3 interacts with eIF4G, this 43S complex is recruited to the 5' end of mRNAs, where it scans the 5'UTR until it finds an AUG start codon within a certain sequence context, marking the beginning of an open reading frame (ORF). At the start codon, the small ribosomal subunit recruits the large ribosomal subunit to form a complete 80S ribosome capable of initiating translation. Translation then proceeds in an elongation cycle, where the ribosome

incorporates amino acids into a polypeptide chain and translocates from codon to codon, until it reaches a stop codon, where the newly synthesized protein is released and the ribosome disassembles into its subunits in order to initiate a new round of translation.

Regulatory RBPs can either increase or decrease translation efficiency. However, most RBPs that regulate translation were found to interfere with translation, and specifically with the initiation step, possibly because initiation is the rate-limiting step of translation or because it is dependent on a large number of proteins that can be targeted (Szostak and Gebauer, 2013). Thus, for translational repression of target mRNAs, RBPs usually use one of many different strategies to interfere with protein-protein or protein-RNA interactions necessary for efficient translation initiation.

1.1.3 RBPs usually bind to the 3' untranslated regions of mRNAs

The major landing site for RBPs to regulate target mRNA stability or translation is the 3' untranslated region (3'UTR), i.e. the mRNA stretch between the stop codon and the poly(A) tail. One reason the 3'UTRs have evolved as repositories for regulatory elements attracting RBPs could be that they are the only part of the mRNA that is free of scanning 43S complexes or translating ribosomes and thus of constraints associated to recognition of ribosomes (Szostak and Gebauer, 2013). At first sight, it may seem inefficient for an RBP to bind to the 3'UTR of an mRNA to regulate a process like translation initiation happening at the 5'end of the transcript. However, in a translation-competent state, mRNAs are thought to circularize due to eIF4G interacting with PABP (Figure 1). This closed-loop formation may allow ribosomes to re-initiate after termination, increasing translational efficiency. It is conceivable that this conformation also allows RNA-binding proteins to more easily interfere with translation initiation when bound to the 3'UTR.

1.2 Mechanisms of RNA-binding proteins to regulate mRNA translation and/or degradation

It is only with the recent advent of different methods for globally identifying RBPs that research starts to appreciate the diversity and complexity of the RBPome in eukaryotic cells (Baltz et al., 2012; Castello et al., 2012; Mitchell et al., 2013). For example, mouse embryonic stem cells were found to express at least 555 different RBPs (Kwon et al., 2013). Taking into account that each individual mRNA binds to a different set of RBPs, the resulting protein production from each mRNA may be a consequence of the activities of all its bound RBPs, sometimes cooperating together or inhibiting each other. Therefore, cells can regulate

specific mRNAs at the post-transcriptional level through elaborate and complex mechanisms. On the other hand, since each RBP has its own RNA-binding specificity and mediates different protein-protein interactions, RBPs have the capacity of shaping gene expression profiles with a huge variety of possibilities. Although RBP research is far from a complete understanding of how the RBPome modulates gene expression, some general principles have emerged by studying the mechanisms of action of individual RBP, some of which are discussed in this section. Considering that the focus of the work presented in this thesis lies on LIN41, an RBP that acts in both translational repression and mRNA degradation by yet unknown mechanisms, this section describes the mechanisms that other RBPs apply to modulate translation and/or mRNA degradation.

1.2.1 The mechanisms of miRNA-associated AGO proteins in post-transcriptional gene regulation

AGO proteins associate with mature miRNAs, untranslated RNA molecules of about 22 nucleotides, to repress protein production on their target mRNAs. To do so, the miRNAs guide AGO proteins to target mRNAs by base-pairing to partially complementary sites, usually in the 3'UTRs. There are several hundred miRNA genes in a typical animal genome and each miRNA is predicted to target many, possibly hundreds of mRNAs. Accordingly, miRNAs are involved in almost every biological process investigated so far, including for instance cell proliferation and differentiation (Shenoy and Blelloch, 2014). They have also been linked to numerous diseases such as cancer and metabolic disorders (Deiuliis, 2016; Reddy, 2015).

On the target mRNAs, miRNAs act as part of a miRNA-induced silencing complex (miRISC) that in its core consists of an AGO protein and a GW182 family protein, recruited by binding to the AGO protein (Figure 1A). GW182 family proteins act as effectors and repress miRNA target mRNAs through translational inhibition, deadenylation and mRNA degradation (Figure 1A). In a current model (Jonas and Izaurralde, 2015), GW182 proteins act as flexible scaffolds to recruit the PAN2-PAN3 complex and the CCR4-NOT complex, with their catalytic subunits responsible for consecutive steps of deadenylation. Deadenylation initiates mRNA degradation through the 5'-to-3' mRNA decay pathway, as for cellular RNAs during regular RNA turnover. Thereby, the CCR4-NOT complex is thought to recruit decapping factors, and decapped mRNAs can be degraded by the exoribonuclease XRN1. The miRISC also inhibits translation of the target mRNAs, but the exact mechanisms that achieve translational repression are still highly debated and an active area of research. Conceivably, deadenylation interferes with the eIF4G-PABP interaction (Figure 1A) and thus opens up the closed-loop mRNA structure, possibly resulting in translational repression. However, recruitment of the CCR4-NOT complex by GW182 proteins is proposed to also induce deadenylation-independent translational inhibition on the miRNA target, possibly through an unknown

mechanism involving the interaction with the putative RNA helicase and translational repressor DDX6 (Chen et al., 2014; Mathys et al., 2014) (Figure 1A). Moreover, different deadenylation- and possibly DDX6-independent mechanisms involving modulation of eIF4A activity and thus inhibition of 43S complex recruitment have also been proposed (Fukao et al., 2014; Fukaya et al., 2014; Meijer et al., 2013) (Figure 1A). In summary, the current models imply that both translational repression and mRNA degradation of miRNA target mRNAs are induced by the same effector proteins and are elicited in parallel, with mRNA degradation as the final fate for a miRNA target. Accordingly, ribosome profiling studies revealed that at steady state, mRNA degradation rather than translational repression is the dominant mechanism of miRNA-mediated target repression (Eichhorn et al., 2014; Guo et al., 2010; Subtelny et al., 2014). Different kinetics of the two mechanisms could explain the observed sequence of events, with translational repression happening shortly before mRNA degradation (Bazzini et al., 2012; Bethune et al., 2012; Djuranovic et al., 2012). However, translational repression of miRNA targets was found to take place without mRNA decay in certain cell types such as early zebrafish embryos (Bazzini et al., 2012), possibly due to deadenylation in the absence of degradative activity in these cells (Subtelny et al., 2014).

1.2.2 The mechanisms of ARE-BPs in modulation of mRNA degradation

Next to AGO proteins, the presumably best-studied RBPs that influence mRNA stability are the members of an RBP class with binding affinity to AU-rich sequence elements (AREs) in the 3'UTR. The mechanisms applied by these ARE-binding proteins (ARE-BPs) are here summarized because they nicely illustrate possible mechanisms for any RBP to regulate mRNA stability. ARE-BPs can have opposite functions in regulating mRNA turnover (Barreau et al., 2005; Hinman and Lou, 2008; Wu and Brewer, 2012). For example, TTP, BRF1 and KSRP seem to induce mRNA degradation on their targets, while HuR seems to protect its targets from degradation. Another ARE-BP, AUF1, can elicit both these mechanisms, possibly due to target-specific or cell-type specific functions. Several models for ARE-BP modes of action have emerged, mainly due to the identification of protein-protein interactions with components of the mRNA decay pathways. Figure 1B depicts an overview of the different interactions of ARE-BPs with other proteins that have been proposed to stabilize or de-stabilize the target mRNA.

To ensure rapid degradation of target mRNAs, ARE-BPs seem to interact with a number of different proteins of the decay pathways. For instance, TTP interacts with the CCR4-NOT complex, the exosome, decapping factors and XRN1 and thus seems to recruit all enzymes needed to degrade the target mRNA through the 5'-to-3' mRNA decay pathway or in a 3'-to-5' manner (Brooks and Blackshear, 2013; Sanduja et al., 2011). BRF1 is a TTP-related protein and likely acts via the same protein-protein interactions

(Sanduja et al., 2011). As another example, KSRP associates with the deadenylase PARN, the decapping factor DCP2 and the exosome, suggesting a similar mechanism with accelerated mRNA decay at both ends of the mRNA (Chou et al., 2006; Gherzi et al., 2004). As a last example, the mechanism for AUF1-mediated decay may involve displacement of PABP from the poly(A) tail and/or recruitment of deadenylases and the exosome (White et al., 2013).

The suggested models for how AUF1 or HuR could stabilize their target mRNAs are rather speculative (Hinman and Lou, 2008; White et al., 2013): Since these proteins bind to AREs that can also be bound by other ARE-BPs, they could protect their targets by occupying the ARE binding site and thereby blocking access for other ARE-BPs. Similarly, binding of AUF1 or HuR could re-structure a neighboring RNA region to inhibit the interaction with RBPs or miRNAs that otherwise would bind to induce degradation. As a third and not mutually exclusive idea, because HuR and AUF1 both have been shown to form oligomers, formation of such oligomers on their target mRNAs could shield a larger region of the RNA from other, destabilizing RBPs.

1.2.3 The mechanisms of RBP in regulation of translation

RBPs that bind to the 3'UTRs of their target mRNAs have evolved diverse modes of translational regulation, mostly by inhibiting translation initiation (Gebauer et al., 2012; Szostak and Gebauer, 2013; Wilkie et al., 2003). To illustrate different mechanisms applied by 3'UTR-bound RBPs to regulate translation, the modes of action of different representative RBPs are discussed here and presented in an overview in Figure 1C. One commonly used strategy is inhibition of the formation of the closed-loop structure, by interfering with the eIF4E-eIF4G or the eIF4G-PABP interaction. RBPs that induce deadenylation usually affect not only mRNA stability, but (in a first step) also translation initiation due to the loss of the eIF4G-PABP interaction. In some environments, such as in oocytes or early embryos of some species, deadenylation can result in pure translational repression, possibly due to a lack of degradative activity at this developmental stage (Subtelny et al., 2014). For example, in both *Drosophila melanogaster* and *Caenorhabditis elegans (C. elegans),* PUF family proteins, together with their associated factors, bind to the 3'UTR and deadenlyate their target mRNAs to repress translation (Miller and Olivas, 2011). In another scenario, RBPs such as *Drosophila* Bruno or *Xenopus* CPEB bind to the 3'UTRs of their targets to recruit 4E-binding proteins (4E-BPs), which bind to eIF4E and block its interaction to eIF4G, resulting in an inhibition of ribosome recruitment and of the closed-loop structure (Nakamura et al., 2004; Stebbins-Boaz et al., 1999). A slightly different strategy is applied by the two *Drosophila* RBPs Bicoid and Brat, which associate with a specific eIF4E isoform called 4EHP that instead of the usual eIF4E isoform

binds to the cap structure, but cannot efficiently bind to eIF4G (Cho et al., 2006; Cho et al., 2005). A second often-used principle by RBPs is a direct inhibition of 43S complex recruitment. Examples for this mode of repression are the mammalian GAIT complex that disrupts the eIF4G to eIF3 interaction and 3'UTR-bound *Drosophila* Sex-lethal (SXL), interfering with 43S complex recruitment through an unknown mechanism without disrupting the eIF4E-eIF4G-PABP interaction (Gebauer et al., 2012; Kapasi et al., 2007). As yet another way of inhibiting translation initiation, the hnRNPs K and E1 bind to the 3'UTR to block the joining of the large ribosomal subunit to the 43S complex at the AUG start codon (Ostareck et al., 1997). Similar to the mechanisms proposed for AUF1 and HuR, RBPs can modulate translation by changing the binding affinities of other RBPs such as AGO proteins. For example, Pumilio and Dnd1 have been shown to increase and decrease miRNA-mediated silencing on their target 3'UTRs, respectively (Kedde et al., 2007; Kedde et al., 2010). Furthermore, some RBPs such as Bruno can promote oligomerization of mRNAs to form larger particles, shielding the packaged mRNAs from access to the translational machinery. More rarely, RBPs have been reported to inhibit translation elongation. Two examples are PUF family proteins and hnRNP E1, both interacting with the elongation factor eEF1A and inhibiting its GTPase activity or its dissociation, respectively (Friend et al., 2012; Hussey et al., 2011).

1.2.4 Some regulatory RBPs bind to the 5' untranslated regions of mRNAs

The vast majority of RBPs that regulate mRNA decay or translation do so by binding to the 3'UTRs of mRNAs. Nevertheless, there are a few studied cases of RBPs that bind to the 5'UTR to modulate translation and their mechanisms of action are illustrated in Figure 1D. Iron regulatory proteins (IRPs) bound to a stem-loop structure in mRNA 5'UTRs sterically inhibit the recruitment of the 43S complex (Kühn, 2015; Volz, 2008). Two different mechanisms are applied by SXL bound to the 5'UTR, it decreases scanning of 43S complexes and increases translation of an upstream open reading frame (uORF), both mechanisms resulting in reduced translation from the main ORF (Beckmann et al., 2005; Medenbach et al., 2011). Scanning of the 43S complex is also inhibited by PABP, when bound to the 5'UTR of its own mRNA (Bag, 2001). Next to these examples of RBPs inhibiting canonical cap-dependent translation, there are some reports on 5'UTR-bound RBPs modulating translation initiation of non-canonical translation modes (not illustrated in Figure 1D). For example, Hu proteins can bind to the 5'UTR and have been suggested to either enhance or inhibit cap-independent translation initiation from an internal ribosome entry site (IRES) (Galban et al., 2008; Kullmann et al., 2002). In addition, some mRNAs harbor terminal oligopyrimidine (TOP) tracts in their 5'UTRs, which attract RBPs such as PTB, TIAR, TIA-1, CNBP, AUF1 or La, modulating translational efficiency by yet unclear mechanisms (Pichon et al., 2012).

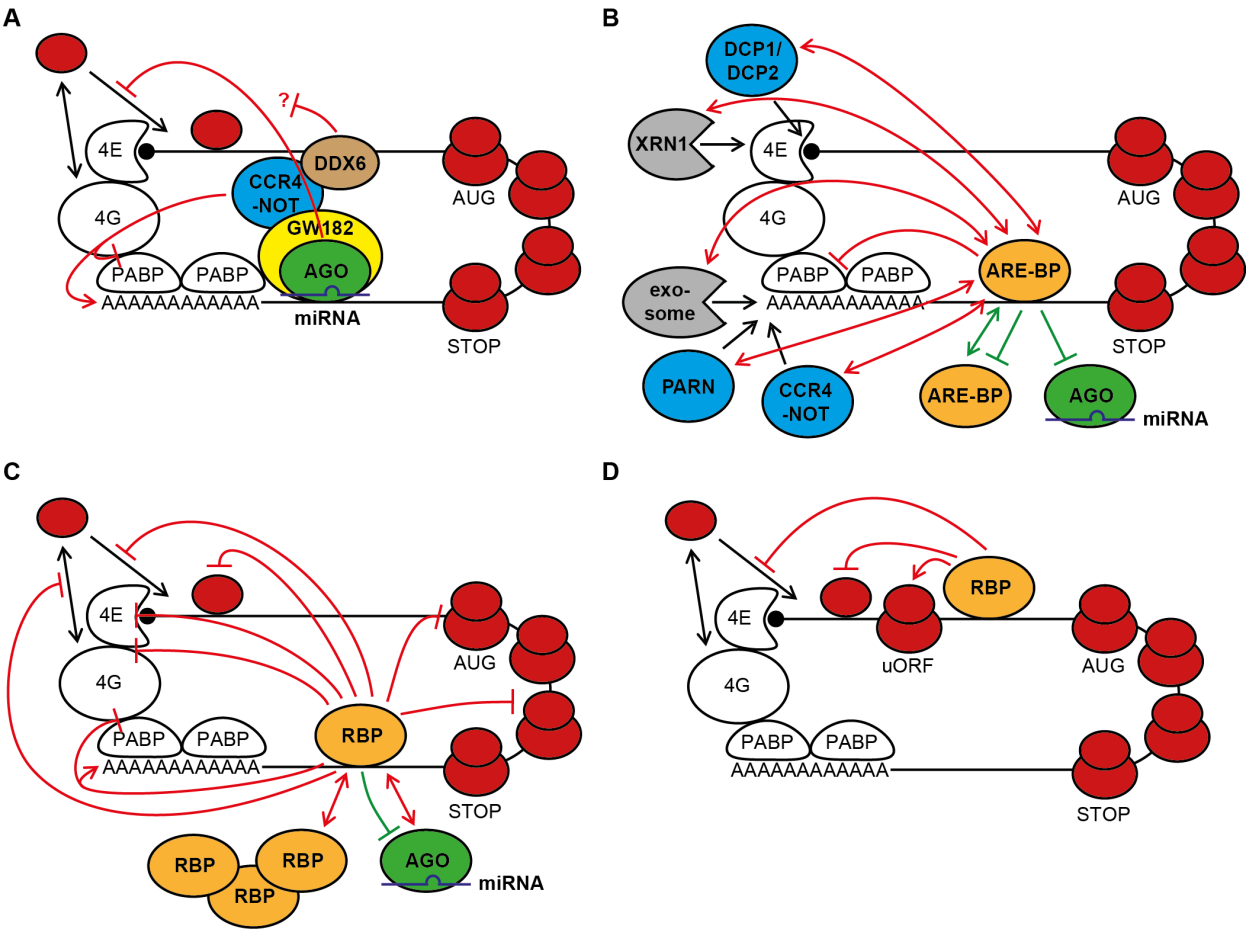


Figure 1. The mechanisms of RBPs to regulate mRNA degradation or translational activity.
(A) Simplified model of the proposed modes of action of miRNAs. Assembly of the miRISC includes the recruitment of deadenylases such as the CCR4-NOT complex. Deadenlyation possibly initiates both translational repression by disrupting the PABP-eIF4G interaction and mRNA degradation through the 5'-to-3' mRNA decay pathway (not drawn). Translational repression may also occur independently of deadenylation, through DDX6 and/or through inhibition of eIF4A helicase activity, leading to unresolved structures in the 5'UTR where the 43S complex cannot land.
(B) Illustration of the different interactions of ARE-BPs to enhance or suppress degradation of target mRNAs. Interactions leading to mRNA degradation are drawn in red, while interactions leading to mRNA stabilization are drawn in green.
(C) Illustration of different possible mechanisms for RBPs bound to the 3'UTR to inhibit translation on their target mRNAs. Interactions or inhibition of interactions that lead to translational repression are drawn in red, the interaction leading to a relief of translational repression is drawn in green.
(D) Illustration of different possible mechanisms for RBPs bound to the 5'UTR to inhibit translation on their target mRNAs.

1.3 RNA-binding proteins in the control of proliferation versus differentiation

Post-transcriptional regulation by RBPs through mechanisms such as those described in section 1.2 is crucial for faithful development of an organism. RBPs regulate many developmental processes, for example clearance of maternal mRNAs in the early embryo, embryonic axes establishment, sex determination or neurogenesis (Kuersten and Goodwin, 2003). Additionally, and most importantly for this thesis, RBPs play an important role in proliferation and differentiation of stem cells and progenitor cells. To date, transcriptional control mechanisms regulating these processes have been much more explored than posttranscriptional control mechanisms, despite evidence that the latter could be equally important (Ye and Blelloch, 2014).

1.3.1 Several RBPs promote stem cell differentiation

At some point in their life, stem and progenitor cells have to switch from a self-renewal to a committed differentiation program in the right location and at the correct time. Several RBPs have been implicated in the control of this switch, including some that control mRNA abundance or translational activity of their targets (Ye and Blelloch, 2014). For instance, the PUF family protein PUM1 promotes exit from the self-renewal state of embryonic stem cells (ESCs) (Leeb et al., 2014). As discussed above, PUF family proteins usually deadenylate their targets upon binding to the 3'UTR, which can lead to either translational repression or mRNA degradation (Miller and Olivas, 2011). Accordingly, PUM1 in ESCs was reported to bind to the 3'UTRs of different transcription factors of the core pluripotency network and thereby decrease their mRNA levels (Leeb et al., 2014). A similar function was proposed for the ARE-BP BRF1, as it seems to bind to AREs in the 3'UTR of core pluripotency transcription factors, resulting in mRNA destabilization (Tan and Elowitz, 2014). A third RBP suggested to promote differentiation of ESCs is ESRP1. It was proposed to associate with mRNAs of core pluripotency transcription factors and to decrease their translation, possibly by binding to their 5'UTRs (Fagoonee et al., 2013).

1.3.2 The RBPs LIN28 and LIN41 promote stem cell proliferation

Opposite to the three above-mentioned RBPs that all promote stem cell differentiation, the two RBPs LIN28 and LIN41 promote the self-renewal state of ESCs (Ye and Blelloch, 2014). In mammals, there are two paralogs of LIN28 referred to as LIN28A and LIN28B. Both have been shown to have similar functions (Rehfeld et al., 2015; Shyh-Chang and Daley, 2013) and are therefore not distinguished in the following. LIN28 and LIN41 are particularly interesting RBPs for stem cell research, because they have been

implicated in self-renewal of not only ESCs, but also of multiple different tissue lineages (Ecsedi and Grosshans, 2013; Rehfeld et al., 2015; Shyh-Chang and Daley, 2013). Yet more strikingly, both LIN28 and LIN41 have been shown to enhance reprogramming efficiency of human fibroblasts into induced pluripotent stem cells (iPSCs) (Worringer et al., 2014; Yu et al., 2007). Although it has remained unclear how exactly LIN28 and LIN41 promote self-renewal, there is an interesting connection between the two RBPs: LIN28 was found to bind to the precursor of the miRNA *let-7* in order to inhibit its maturation (Heo et al., 2008; Newman et al., 2008; Rybak et al., 2008; Viswanathan et al., 2008). On the other hand, LIN41 is a conserved target of the miRNA *let-7* (Ecsedi and Grosshans, 2013), suggesting that LIN28 indirectly promotes expression of LIN41 by decreasing *let-7* levels. As described in section 1.4, these interconnections are conserved in *C. elegans*, where LIN28 also inhibits *let-7* maturation (Lehrbach et al., 2009; Van Wynsberghe et al., 2011), and *let-7* also regulates LIN41 expression (Reinhart et al., 2000; Slack et al., 2000).

1.3.3 Possible mechanisms of LIN28 in promoting self-renewal

LIN28 has been studied extensively since the discovery that it enhances the reprogramming efficiency for generating iPSCs (Yu et al., 2007). Thereby, it has become clear that it not only binds to *let-7* precursor transcripts, but also to thousands of mRNAs, including transcripts encoding cyclins, splicing factors, metabolic enzymes and ribosomal proteins (Balzer et al., 2010; Cho et al., 2012; Hafner et al., 2013; Li et al., 2012; Peng et al., 2011; Polesskaya et al., 2007; Wilbert et al., 2012; Xu et al., 2009). When bound to target mRNAs, LIN28 usually seems to promote their translation. Although many proteins involved in translational regulation, such as eIF3, eIF4E, elongation factors, PABP, IGF2BPS, MSI1, RHA and some ribosomal proteins, have been found to interact with LIN28 (Balzer and Moss, 2007; Jin et al., 2011; Polesskaya et al., 2007), the mechanism of LIN28-mediated translational enhancement has remained unclear. Moreover, it is currently unknown whether LIN28 promotes stem cell proliferation mainly through *let-7*, by promoting translation of its direct mRNA targets, through its additional functions in the nucleus or by a combination of those functions (Shyh-Chang and Daley, 2013). If it mainly acts through its cytoplasmic mRNA targets, it will be important to find which of the thousands of proposed targets are involved in its role in regulation of stem cell self-renewal. This is a very challenging task in a system like ESCs and it was therefore proposed that such an undertaking should rather be ventured in a model system with powerful genetics like *C. elegans* (Shyh-Chang and Daley, 2013).

1.3.4 Possible mechanisms of LIN41 in promoting self-renewal

Compared to LIN28, much less is known about the mechanisms of LIN41 in promoting stem cell self-renewal, possibly due to its more recent implication in increasing the reprogramming efficiency of fibroblasts to generate iPSCs (Worringer et al., 2014). Nevertheless, LIN41 seems to emerge as a highly conserved regulator of self-renewal, differentiation, and cell fate plasticity. For example, it was implicated in control of proliferation versus differentiation of mouse and human embryonic stem cells (Chang et al., 2012; Rybak et al., 2009; Worringer et al., 2014). Moreover, it seems to be important for neuronal development and plasticity, as it stimulates proliferation and inhibits premature differentiation of mouse neural progenitor cells (Chen et al., 2012; Cuevas et al., 2015), and promotes axon regeneration of *C. elegans* neurons (Zou et al., 2013). The defect in proliferation of neural progenitor cells is the likely cause of the embryonic lethality observed in LIN41 mutant mice, which display a defect in neural tube closure (Chen et al., 2012).

Interestingly, mammalian LIN41 seems to have two different molecular functions, consistent with it belonging to the TRIM-NHL protein family. TRIM-NHL proteins like LIN41 consist of an N-terminal TRIM domain (a Tripartite Motif with RING, B-box and coiled-coil domains), characteristic of proteins with E3 ubiquitin ligase activity (Ikeda and Inoue, 2012), and a C-terminal NHL (NCL-1, HT2A2, and LIN-41) repeat domain, which may mediate sequence-specific RNA binding (Loedige et al., 2015; Loedige et al., 2014). Accordingly, mammalian LIN41 has been shown to both ubiquitylate target proteins (Chen et al., 2012; Rybak et al., 2009) and to repress target mRNAs (Chang et al., 2012; Loedige et al., 2013; Worringer et al., 2014), and both mechanisms have been proposed to be involved in promoting stem or progenitor cell proliferation (Chang et al., 2012; Chen et al., 2013; Rybak et al., 2009; Worringer et al., 2014). However, as the E3 ubiquitin ligase activity does not seem to be conserved in *C. elegans* (Tocchini et al., 2014), where LIN41 also regulates self-renewal (see section 1.4), it is more likely that the conserved function of LIN41 in controlling self-renewal activity is due to its function as an RBP.

As an RBP, LIN41 was proposed to repress its mRNA targets by binding to their 3'UTRs to induce both degradation and translational repression. Although in contrast to LIN28, there are no reported studies with RNA co-immunoprecipitations coupled to RNA sequencing for LIN41, some individual mRNAs have been shown to be bound by LIN41 (Chang et al., 2012; Loedige et al., 2013; Mitschka et al., 2015; Worringer et al., 2014). All these mRNAs were observed to be reduced in abundance in the presence of LIN41, establishing LIN41 as an RBP that mediates mRNA degradation. However, there are at least two indications that LIN41 could also act as a translational repressor. First, some reporters with LIN41 target 3'UTRs were changed more extensively on the protein level than on the RNA level (Loedige et al., 2013), similar to what has been observed with miRNA target reporters. Second, as described in section 1.4, a

potential target of LIN41 in *C. elegans*, *lin-29*, has been observed to be upregulated specifically on the protein level at the time in development during which LIN41 is downregulated (Bettinger et al., 1996; Rougvie and Ambros, 1995; Slack et al., 2000). However, as *lin-29* has not been shown to be a direct LIN41 target (see section 1.4), a LIN41-mediated repression of *lin-29* is pure speculation. Therefore, based on the data from Loedige *et al*. (Loedige et al., 2013), a current model for the LIN41 mechanism on target mRNAs could be that it acts similar to a miRNA-Ago complex (section 1.2), inducing translational repression and mRNA degradation in parallel when bound to the 3'UTR of a target mRNA.

As for LIN28, it is unknown which targets are relevant for the function of LIN41 in promoting stem cell self-renewal. In order to find such physiologically relevant targets, experiments are preferentially performed *in vivo* rather than in cell culture. As the role of LIN41 in stem cell proliferation is conserved in *C. elegans*, we sought to study the nematode homolog of LIN41 in this thesis. In *C. elegans*, LIN41 regulates the proliferation of stem cell-like cells in the skin, as part of a pathway called the heterochronic pathway that is introduced in section 1.4.

1.4 *Caenorhabditis elegans* and the heterochronic pathway

1.4.1 Caenorhabditis elegans and its life cycle

Caenorhabditis elegans (*C. elegans*) is a free-living, about one millimeter long roundworm that lives in temperate soils of the whole world and feeds on bacteria and other microorganisms. Several features make it an exceptional model organism, including the following: i) There are two sexes, hermaphrodites and males. Whereas hermaphrodites are self-fertilizing and therefore easy to expand to a whole population of worms with identical genotype, males can be used to introduce and combine mutations. ii) *C. elegans* is transparent, allowing for visualization of all cells in all tissues using differential interference contrast (DIC) optics and for tracking protein localization and dynamics *in vivo* with the help of fluorescent proteins. iii) It has very powerful genetics, with well-described phenotypes that can be used as markers, genetic balancers to allow for maintenance of lethal or sterile mutant alleles, established tools for straightforward integration of single-copy or multi-copy transgenes and a rapidly growing CRISPR-Cas9 toolkit for targeted genome editing. iv) It has a rapid life cycle of about 3 days at 25 °C (Figure 2A), allowing for efficient crossings, rapid expansion of worm populations and straightforward experiments to study developmental biology (see also section 1.4.2).

The life cycle of *C. elegans* consists of embryogenesis, four larval stages named L1, L2, L3 and L4, and an adult stage, during which sexually mature hermaphrodites lay about 300 eggs (Figure 2A). Hermaphrodites can self-fertilize their oocytes because they produce their own sperm, or, after mating, can use the male sperm for fertilization. The initial steps in embryogenesis occur in the uterus of the mother, before the embryos are laid at about the 30-cell stage and continue their development outside the mother. The first larval stage in postembryonic development, the L1 stage, starts when the worm hatches and begins feeding. Every larval stage ends with a molt, during which a new cuticle is produced and the old cuticle is shed. With the fourth and final molt, the larvae turn into adults. This larval-to-adult transition is a highly regulated event (see below), as it comprises the beginning of a developmental program distinct from the larval program, for instance ensuring that molting will never occur again, that an adult culticle with specialized features is synthesized and that the sexual organs mature.

1.4.2 The heterochronic pathway controls the timing of stage-specific developmental events

In order to study how developmental events are regulated, *C. elegans* is an extremely powerful model organism, because all somatic cells (959 in the hermaphrodite, 1031 in the male) can be tracked during development. Therefore, cell fate decisions for each cell have been observed and described (Kimble and Hirsh, 1979; Sulston and Horvitz, 1977). These studies have revealed that the cell divisions occurring over development were largely invariant between individual worms, i.e. both the identity of the daughter cells and the timing of the division were the same in each animal. Taking advantage of this invariance, research with *C. elegans* mutants led to the discovery of many genes with roles in controlling developmental fates, including genes that, when mutated, changed the time point of developmental events. These mutations, unlike others, do not cause an overall slower or faster developmental rate, but instead result in a changed timing of a specific developmental event relative to other developmental events in the same organism. Mutants with such a phenotype were therefore called heterochronic (Greek: heteros = other, chronos = time) (Ambros and Horvitz, 1984). In these animals, specific cells either prematurely adopted fates of later developmental stages or repeated fates of earlier developmental stages. Because a few years later, the four different heterochronic genes known back then were found to suppress or enhance each other's phenotypes, they were proposed to be members of the same pathway (Ambros, 1989), later named the heterochronic pathway. Up to now, many more heterochronic genes have been identified and their position in the pathway determined by genetic criteria. Those with the strongest and clearest phenotypes make up the core pathway (Rougvie and Moss, 2013) and regulate the timing of events specific to a certain larval stage (see below). Although the heterochronic pathway regulates the developmental timing of cell

divisions in different tissues, much of its research has focused on the particularly interesting and easy-to-observe cell divisions of the epidermal seam cells in the worm skin.

1.4.3 Heterochronic mutants change the timing of seam cell divisions and differentiation

Seam cells are stem-cell like cells, aligned symmetrically on each side of the worm body. Most seam cells go through an asymmetric cell division during each of the first three larval molts (Figure 2B). These cell divisions give rise to another seam cell and a cell that ceases proliferation (Sulston and Horvitz, 1977). The latter cell differentiates and fuses with an epidermal syncytium called hyp7. Almost the whole body of the worm is covered by hyp7, and this large single cell grows more in each larval stage, because it acquires additional nuclei through fusion events with other cells. At the last molt, seam cells do not divide but instead terminally differentiate by fusing to each other to form a syncytium (Figure 2B,C). The seam syncytium is thought to be important for synthesis of an adult cuticle and specifically for synthesis of ridges within the adult cuticle called alae (Figure 2C), extending over the whole length of the cuticle (Singh and Sulston, 1978). Because alae are not present on the L4 stage cuticle, their presence is often used as a readout for terminal differentiation of the seam cells and for an adult epidermis in general. In summary, at the larval-to-adult transition, the cells of the worm epidermis go through a transition from juvenile to adult cell fates. This larval to adult switch (L/A switch) was originally defined to include four developmental events that can be easily observed (Ambros, 1989): Termination of seam cell divisions, seam cell fusion to a syncytium, formation of an adult cuticle and exit from the molting cycle (Figure 2C). In heterochronic mutants, L/A switch events either occur too early, too late or never. Hence, the heterochronic pathway ultimately regulates the transition from juvenile to adult epidermal cell fates.

Prior to the L/A switch, the different seam cells go through slightly different fates. The 10 pairs of seam cells are named H0-H2, V1-V6 and T (from head to tail), in the order they appear in hatched L1 larvae (Figure 3A). With the exception of H0, all seam cells undergo stem-cell like divisions during the larval stages. However, the heterochronic pathway is usually studied by observing only the V1-V4 and V6 cells (bold in Figure 3A), because those go through identical cell divisions during development. The V1-V4 and V6 cells asymmetrically divide during each larval stage (Figure 3B,C,D), but prior to the asymmetric division of the L2 stage, they symmetrically divide to produce two undifferentiated seam cells (Figure 3B,C,D). This symmetric division gives rise to a pair of seam cells of the same origin, therefore labeled with identical names in Figure 3B. At the larval-to-adult transition, after going through the last asymmetric division in the early L4 stage, the V1-V4 and V6 cells form a syncytium through fusion with each other and all other seam cells, including those of the H and T lineage (16 seam cells fuse in total).

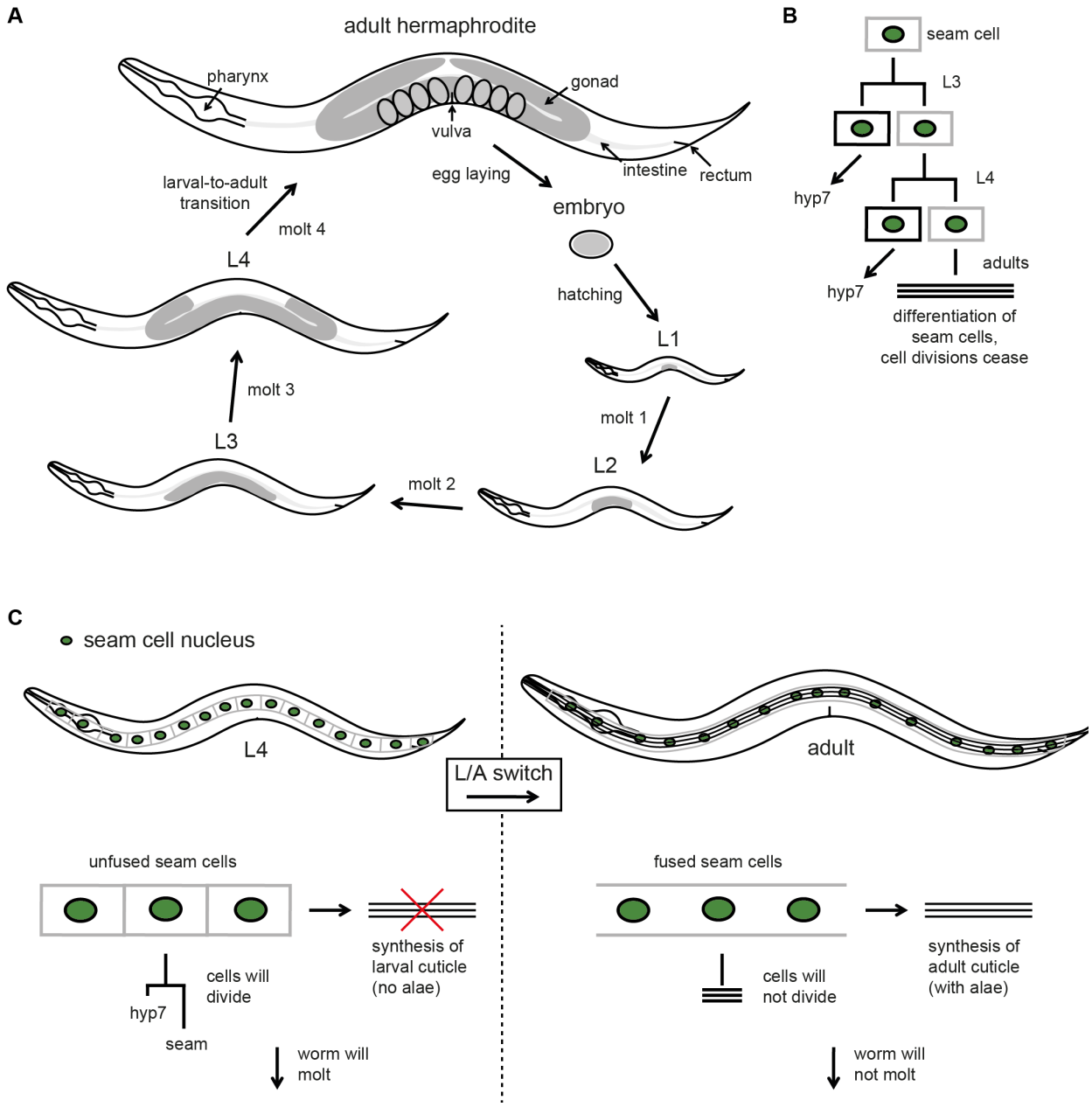


Figure 2. The developmental transition from larva to adult in *C. elegans*
(A) The life cycle of *C. elegans*. After hatching, worms develop through four larval stages (L1-L4), each followed by a molt. At the final molt, they develop into adult worms that sexually mature. Adult hermaphrodites lay eggs, from which the next generation of worms will hatch.
(B) The asymmetric seam cell division. At the end of larval stages L1-L3, at about the time of the molt, seam cells divide and give rise to another seam cell (gray) and a cell that fuses to the hyp7 syncytium (black). By contrast, at the L4-to-adult molt, the seam cells have terminally differentiated and do not divide again. Seam cell nuclei can be observed with a specific GFP marker (section 2.6) and are therefore drawn in green.
(C) The L/A switch consists of four distinct developmental events: Fusion of the seam cells, cessation of seam cell divisions, cessation of the molting cycle and synthesis of an adult cuticle with alae.

1.4.4 Precocious versus retarded phenotypes

Two different kinds of epidermal phenotypes are observed for heterochronic mutants. On the one hand, mutations in genes such as *lin-14*, *lin-28*, *lin-41* or *hbl-1* result in a precocious phenotype, i.e. stage-specific events are skipped and L/A switch events occur at an earlier molt (Figure 3C). On the other hand, mutations in genes like *lin-4*, *let-7*, the three *let-7* sisters or *lin-29* cause a retarded phenotype, i.e. stage-specific events are repeated and the L/A switch fails to occur, resulting in "adults" with a larval cuticle (Figure 3D). Thus, in adults of retarded mutants, additional seam cell divisions and extra molts can be observed, while alae are not produced, corresponding to a repetition of the larval program in the epidermis (Figure 3D). All discovered heterochronic genes of the core pathway are either transcription factors, RNA-binding proteins or miRNAs. Thus, the pathway seems to take advantage of both transcriptional and post-transcriptional mechanisms. However, the direct targets of the heterochronic transcription factors LIN-14, HBL-1 and LIN-29 are largely unknown, and the same is true for those of the RNA-binding proteins LIN28 and LIN41. Therefore, the direct molecular links between the different heterochronic genes are mostly unclear, although from genetic experiments, the genes have been positioned relative to each other (Figure 3E). The described links between the heterochronic miRNAs and their targets, which are down-regulated by binding of the miRNAs to their 3'UTRs, are thus the clearest molecular links of the pathway known to date (see below). In the following sections, the heterochronic genes of the core pathway are discussed. The roles of other heterochronic genes such as *daf-12, lin-42* or *lin-46* are less important for this thesis and thus not explained, but are described in several reviews (Moss, 2007; Rougvie and Moss, 2013).

1.4.5 The heterochronic transcription factors regulate developmental switches

Three transcription factors, LIN-14, HBL-1 and LIN-29, are part of the core heterochronic pathway. In general, they are the key regulators of developmental switches and their expression is regulated by more upstream heterochronic genes encoding for miRNAs and RBPs. Already in 1989, when Victor Ambros first described the heterochronic pathway, *lin-29* was found to be the most downstream gene and thus the ultimate target of the pathway (Figure 3E) (Ambros, 1989). Although many more heterochronic genes were discovered up to now, this still seems to hold true. Animals mutated for *lin-29* never go through the L/A switch and infinitely repeat the larval cell fates in the hypodermis (Ambros, 1984) (Figure 3D). In the epidermis of wild-type worms, LIN-29 was found to be expressed only in L4 stage worms, after the final seam cell division and before the last molt and the L/A switch occur (Bettinger, 1996). Accordingly, *lin-29* mutants normally progress through all larval stages until they reach the end of the L4 stage, where their epidermis fails to become adult and instead goes through the program of the L4 stage again (Ambros,

1984) (Figure 3D). Thus, it was concluded that the crucial event for the L/A switch to occur is the upregulation of LIN-29. In all heterochronic mutants where the L/A switch occurs prematurely, the phenotype can be suppressed by additionally mutating *lin-29*, leading to repeated larval cell fates as in *lin-29* single mutants (Rougvie and Moss, 2013). Hence, it is likely that the premature L/A switch in every precocious heterochronic mutant is the result of premature upregulation of the LIN-29 protein in the epidermis. In conclusion, the key mechanism to time the L/A switch seems to be restriction of LIN-29 activity in earlier larval stages through the heterochronic genes. However, it is still unknown which mechanisms regulate LIN-29 activity at the molecular level (Rougvie and Moss, 2013).

The two other transcription factors of the core heterochronic pathway, LIN-14 and HBL-1, seem to regulate developmental transitions in a manner opposite to that of LIN-29. Just like LIN-29 needs to be upregulated for the transition to another developmental stage, LIN-14 and HBL-1 need to be downregulated for such a transition to occur (Figure 3F). A downregulation of LIN-14 is crucial for the L1-to-L2 transition (Ambros and Horvitz, 1987; Lee et al., 1993; Wightman et al., 1993), while a downregulation of HBL-1 seems to promote the L2-to-L3 transition (Abbott et al., 2005; Abrahante et al., 2003; Lin et al., 2003) (Figure 3E). To date, it is not clear if and how the L3-to-L4 transition is regulated by the heterochronic pathway (Rougvie and Moss, 2013), although LIN41 was suggested to play a role in this transition ((Vadla et al., 2012), and see below). It is very clear that the transcription factor LIN-14 promotes the gene expression program that keeps the worms in the L1 stage, while a downregulation of LIN-14 is sufficient to induce the transition to the L2 stage. Thus, *lin-14* loss-of-function (lf) mutants skip the L1 stage and directly start the L2 stage specific patterns, with a symmetric seam cell division in the L1 stage (Ambros and Horvitz, 1987) (Figure 3C). As the following larval stage programs also take place one stage too early, *lin-14(lf)* mutant animals go through the L/A switch at the L3-to-L4 molt (Figure 3C). On the other hand, *lin-14* gain-of-function (gf) mutants repeat the L1 stage seam cell patterns forever and never transit to an L2-specific program nor ever reach the L/A switch (Ambros and Horvitz, 1984) (Figure 3D).

For which developmental transition(s) HBL-1 activity has to be downregulated is not entirely clear. The main problem is that our knowledge is based on hypomorphic reduction-of-function (rf) *hbl-1* alleles, as a full depletion of HBL-1 leads to embryonic lethality (Abrahante et al., 2003; Lin et al., 2003). These rf mutant animals show signs of a precocious L/A switch at the L3-to-L4 molt, such as alae, fusion of seam cells and precocious LIN-29 expression, suggesting they skip one of the larval stage programs (Abrahante et al., 2003; Lin et al., 2003). However, because it is unclear, which larval stage *hbl-1(rf)* mutant animals would skip, and because they showed normal symmetric seam cell divisions at the L2 stage, their seam cell division patterns are usually drawn as normally occurring up to the L3-to-L4 molt, where they differentiate (alae are visible) (Figure 3C). In the L4 stage, additional seam cell divisions were observed,

suggesting that the complete L/A switch program thus only occurred at the larval-to-adult transition (Abrahante et al., 2003; Lin et al., 2003) (Figure 3C). Nevertheless, there are some indications that these complicated phenotypes may be a consequence of the partial HBL-1 depletion and that a full depletion of HBL-1 activity would simply lead to skipping of the L2 stage pattern (Figure 3C): First, depletion of *hbl-1* by RNAi leads to fewer seam cells at the L3 stage, suggesting that the symmetric division was skipped (Abrahante et al., 2003). Second, as described below, mutation of all three *let-7* miRNA sisters leads to repetition of the L2 stage pattern, a phenotype that can be suppressed by depletion of HBL-1 (Abbott et al., 2005). Third, in the vulva region, L3-stage specific cell division patterns are observed one stage too early, on *hbl-1* RNAi but also in *hbl-1(rf)* mutant animals (Abrahante et al., 2003). In summary, it seems likely that HBL-1 activity needs to be downregulated to allow progression through the L2-to-L3 transition (Figure 3E).

1.4.6 The miRNAs of the heterochronic pathway

Out of the ten core genes of the heterochronic pathway, five are miRNAs. These include *lin-4*, the first discovered miRNA (Lee et al., 1993), *let-7*, whose sequence is perfectly conserved in mammals (Pasquinelli et al., 2000; Reinhart et al., 2000), and three *let-7* sisters, which have the same seed sequence as *let-7*, but act earlier in development than *let-7* (Abbott et al., 2005). Mature miRNAs of *lin-4, let-7* or the group of three *let-7* sisters are upregulated at distinct times in larval development and thereby initiate the progression of worm tissues such as the epidermis through the larval stages (Figure 3F). Generally, animals mutant for these miRNAs show retarded phenotypes, because one of the larval stage transitions cannot occur (Figure 3D).

The L1-to-L2 switch is controlled by the miRNA *lin-4*, which is upregulated in the L1 stage (Figure 3F) and downregulates the transcription factor LIN-14 and the RBP LIN28 by binding to the 3'UTR of their mRNAs (Lee et al., 1993; Moss et al., 1997; Wightman et al., 1993). As described above, downregulation of LIN-14 is crucial for the L1-to-L2 switch, whereas the role of LIN28 is less clear ((Rougvie and Moss, 2013), and see below). Due to the failure in LIN-14 downregulation, *lin-4(lf)* mutant animals show the identical phenotype as those with a *lin-14(gf)* mutation, i.e. they repeat the L1 stage-specific pattern of seam cell division and never go through the symmetric division characteristic for the L2 stage (Ambros and Horvitz, 1984) (Figure 3D). From these results, and because LIN-28 seems to act only later in development, it is clear that the L1-to-L2 switch is regulated by *lin-4* specifically regulating its key target *lin-14.* Consistent with this, *lin-4(lf); lin-14(lf)* double mutants suppress the *lin-4(lf)* mutant phenotype and instead show the *lin-14(lf)* phenotype, i.e. they are precocious mutants due to skipping of the L1 stage pattern.

The L2-to-L3 switch is initiated by upregulation of the three *let-7* sisters *mir-48*, *mir-84* and *mir-241* (Figure 3F) (Abbott et al., 2005). These three miRNAs share the same seed sequence and are upregulated at the same developmental time to redundantly act on their main target, the transcription factor HBL-1 (Abbott et al., 2005). Triple mutants for *mir-48*, *mir-84* and *mir-241* repeat the L2 stage-specific symmetrical seam cell division, resulting in extra seam cells at the L3 stage (Figure 3D). If in later development, the L2 stage-specific pattern is repeated again or if the animals progress through further larval stage patterns is yet unclear (Figure 3D). Although other heterochronic genes such as *lin-41* and *lin-28* were found to have *let-7* family complementary sites in their 3'UTRs, *hbl-1* was found to be the main target of the *let-7* sisters. This was because *lin-41* did not suppress the *mir-48*; *mir-84; mir-241* triple mutant phenotype, and because *hbl-1*, but not *lin-28*, was downregulated dependent on the three *let-7* sisters and on its 3'UTR (Abbott et al., 2005).

Finally, upregulation of the *let-7* miRNA during the L3 and L4 stages (Figure 3F) initiates the transition from the L4 stage to adults (Reinhart et al., 2000). Thus, in *let-7(lf)* mutant animals, seam cell terminal differentiation and exit from the cell cycle fail, leading to a repetition of the L4-stage program (Figure 3D). Eventually, *let-7* regulates the upregulation of the transcription factor LIN-29, as LIN-29 was not detected in L4 stage *let-7* mutant animals (Reinhart et al., 2000). Because LIN-29 needs to be upregulated for the L/A switch to occur, *let-7* has to act indirectly on LIN-29, by repressing a repressor of LIN-29 expression. At the time when *let-7* was discovered, the main *let-7* target was proposed to be LIN41 (Reinhart et al., 2000; Slack et al., 2000), but later, other *let-7* targets such as HBL-1 (Abrahante et al., 2003; Lin et al., 2003) and DAF-12 (Grosshans et al., 2005) were discovered. Surprisingly, individual depletion of all these targets resulted in suppression of *let-7* phenotypes, restoring events of the larval-to-adult transition such as alae secretion (Abrahante et al., 2003; Grosshans et al., 2005; Lin et al., 2003; Slack et al., 2000). It has thus remained unclear which *let-7* targets have to be downregulated in order for the larval-to-adult transition to occur. Because LIN41 has recently been shown to be the main target for the function of *let-7* in vulval development (Ecsedi et al., 2015), and because LIN41 is downregulated during the L3 and L4 stages, at the time of the main boost in *let-7* upregulation (Figure 3F), it is likely that LIN41 is one of the key targets of *let-7* for its function in the larval-to-adult transition (Figure 3E).

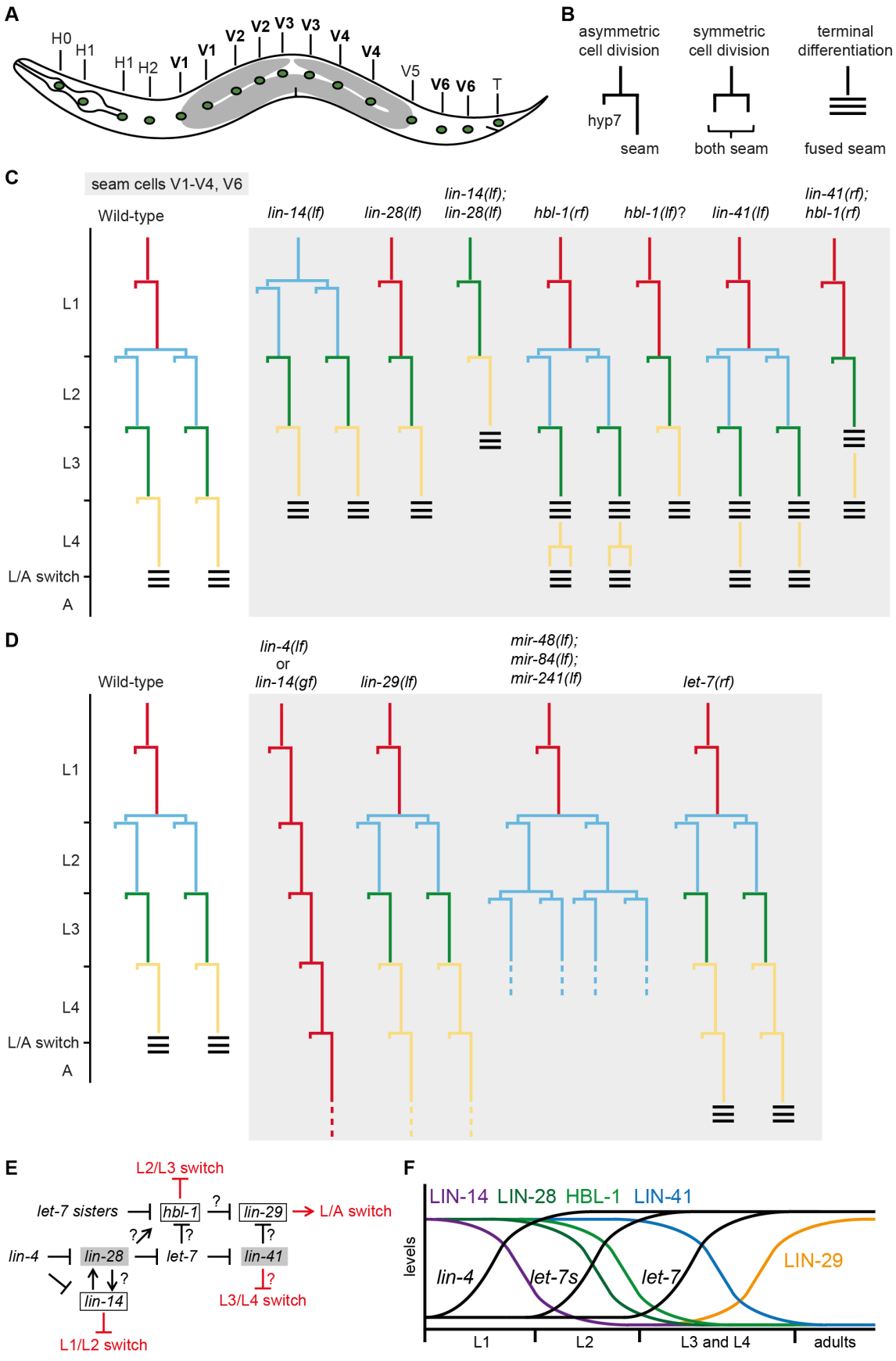

A
H0
H1
H1
H2
V1
V1
V2
V2
V3
V3
V4
V4
V5
V6
V6
T
B
asymmetric cell division
symmetric cell division
terminal differentiation
hyp7
seam
both seam
fused seam
C
seam cells V1-V4, V6
Wild-type
lin-14(lf)
lin-28(lf)
lin-14(lf); lin-28(lf)
hbl-1(rf)
hbl-1(lf)?
lin-41(lf)
lin-41(rf); hbl-1(rf)
L1
L2
L3
L4
L/A switch
A
D
Wild-type
lin-4(lf) or lin-14(gf)
lin-29(lf)
mir-48(lf); mir-84(lf); mir-241(lf)
let-7(rf)
L1
L2
L3
L4
L/A switch
A
E
L2/L3 switch
let-7 sisters
hbl-1
lin-29
L/A switch
lin-4
lin-28
let-7
lin-41
lin-14
L3/L4 switch
L1/L2 switch
F
LIN-14 LIN-28 HBL-1 LIN-41
levels
lin-4
let-7s
let-7
LIN-29
L1
L2
L3 and L4
adults

Figure 3: Precocious and retarded phenotypes in mutants of heterochronic genes
(A) The seam cells in a wild-type adult hermaphrodite.
(B) Schematic depiction of asymmetric and symmetric cell divisions as well as terminal differentiation events.
(C) Seam cell lineage patterns for precocious mutants of the core heterochronic pathway.
(D) Seam cell lineage patterns for retarded mutants of the core heterochronic pathway.
(E) Model for the heterochronic pathway regulating the switches between the different developmental stages. The proposed regulations of the developmental switches are colored in red, with the link between *lin-41* and the L3/L4 switch being very speculative (red question mark), while unclear molecular links are labelled with a black question mark.
(F) Schematic epidermal gene expression profiles of the heterochronic miRNAs, RBPs and transcription factors. The upregulation of mature miRNAs are depicted with black profiles, while the down- or upregulation of transcription factors or RBPs are drawn with different colors.

1.4.7 The RBPs of the heterochronic pathway

Compared to the roles of the transcription factors and of the miRNAs, the roles of LIN28 and LIN41, the two RBPs in the heterochronic pathway, have remained more mysterious. Through genetic experiments, the RBPs are placed in between miRNA genes and transcription factors of the heterochronic pathway. Accordingly, they seem to be miRNA targets and to directly or indirectly regulate the expression of transcription factors (Figure 3E). LIN28 is targeted already early on in development by the *lin-4* miRNA (Moss et al., 1997), but it seems to have roles in different switches between larval stages. In its best-studied function, LIN28 inhibits the expression of the mature *let-7* miRNA, preventing its premature accumulation in early larval stages (Figure 3F): On the one hand, LIN28 binds to the precursor of the *let-7* miRNA to inhibit its processing to the mature form, at the same time targeting it for degradation through recruitment of a poly(U) polymerase (Lehrbach et al., 2009; Viswanathan et al., 2008). On the other hand, LIN28 binds to the primary *let-7* transcript in the nucleus, preventing its processing to pre-*let-7* by Drosha (Van Wynsberghe et al., 2011). From these functions of LIN28, it could be expected that due to premature upregulation of *let-7*, *lin-28* mutants go through a precocious larval-to-adult transition one stage too early, without any defects in earlier larval stages. However, *lin-28* mutants specifically skip the L2 stage pattern of seam cell divisions (Figure 3C) (Ambros and Horvitz, 1984), independent of premature *let-7* accumulation (Vadla et al., 2012). This suggests that LIN28 has an additional function in the heterochronic pathway, namely promotion of L2 stage events. LIN28 positively regulates the expression of HBL-1 (Vadla et al., 2012), which could explain this additional function, but how this could work on the mechanistic level has remained unclear. Adding another layer of complexity, *lin-28* also genetically interacts with *lin-14*. In *lin-28(lf)* mutants, *lin-14* is repressed, and in *lin-14(lf)* mutants, *lin-28* is repressed (Arasu et al., 1991; Moss et al., 1997). Hence, LIN-14 and LIN28 promote each other's expression, but again by unknown mechanisms (Figure 3E).

The other RBP of the core heterochronic pathway is the TRIM-NHL protein LIN41. Due to its domain architecture, described in section 1.3, it has been speculated to act as an RBP or as an E3 ubiquitin ligase in the pathway. Expression of the LIN41 is regulated by the miRNA *let-7*, targeting the 3'UTR of the *lin-41* mRNA to downregulate its expression during the L3 and L4 stages (Reinhart et al., 2000; Slack et al., 2000). The exact role of LIN41 in the heterochronic pathway has been unclear, but its downregulation by *let-7* has been proposed to be important for both the L3-to-L4 transition and the L4-to-adult transition (Slack et al., 2000; Vadla et al., 2012) (Figure 3C). In *lin-41(lf)* mutant animals, the seam cells differentiate one stage too early, as observed by precocious alae secretion (Slack et al., 2000) (Figure 3C). However, this phenotype occurs only in about half the animals and only for a subset of seam cells, with the other animals and all other seams cells differentiating normally at the larval-to-adult transition (Figure 3C). Interestingly, LIN41 overexpression results in a similar retarded phenotype as for *let-7(lf)* animals, with a repetition of the larval stage pattern of seam cell divisions (Slack et al., 2000) (Figure 3D). Therefore, LIN41 overexpression was suggested to be the reason for the *let-7(lf)* mutant phenotype, possibly acting by repressing LIN-29 expression (Slack et al., 2000). Because the *lin-29* mRNA was detected two larval stages earlier than the LIN-29 protein (Bettinger et al., 1996; Rougvie and Ambros, 1995), it was proposed that LIN41 may inhibit the translation of the *lin-29* mRNA or target the LIN-29 protein for degradation through ubiquitylation (Slack et al., 2000). However, none of these models has been confirmed to date, and LIN41 has remained a mysterious protein that acts in the heterochronic pathway to regulate unknown targets through unknown mechanisms.

1.4.8 Extensive redundancy in the heterochronic pathway

As described above, *lin-4* or *lin-29* mutants have fully penetrant retarded phenotypes (they never go through the L/A switch) and thus do not seem to share redundancy with other heterochronic genes (Ambros and Horvitz, 1984). By contrast, genetic evidence suggests that all other core factors of the heterochronic pathway have redundant functions with other genes. The *let-7* sisters *mir-48*, *mir-84* and *mir-241* seem to act redundantly in target repression, as single mutants and double mutant combinations only lead to partially penetrant phenotypes when compared to the triple mutant (Abbott et al., 2005). A hint to the redundancy between the *let-7* sisters could be their similar temporal expression patterns, potentially leading to repression of the same target(s) through their common seed sequence. In this model, possible explanations for the redundancy are that either the expression levels of each sister miRNA are not high enough for full target repression, or that the three miRNAs act in different tissues. Redundancy has also been suggested for the miRNA *let-7*, because *let-7(rf)* mutants execute only a single extra larval molt and seam cell division cycle, before they terminally differentiate (Reinhart et al., 2000)

(Figure 3D). Therefore, additional, *let-7* independent mechanisms could also regulate the expression of LIN-29. However, since *let-7(lf)* mutants die as young adult animals due to a defect in vulval development (section 2.5), these conclusions are based on the phenotypes of animals with a partial loss of *let-7* activity. Therefore, it remains to be determined whether *let-7* indeed regulates LIN-29 redundantly with some other factor or whether a full loss of *let-7* activity in the epidermis would lead to phenotypes identical to those of *lin-29(lf)* mutants.

Clear redundancies are observed for all core heterochronic genes leading to precocious phenotypes. While *lin-14* and *lin-28* single mutants go through the L/A switch one stage too early, the double mutant does so two stages too early (Ambros, 1989) (Figure 3C). The mechanisms underlying the enhancement of this precocious phenotype are not yet known and neither is which larval stages are skipped in these double mutant animals (in Figure 3C, the depicted pattern skips the L1 and L2 stages, but this is just a speculation). The second evident redundancy is that of *lin-41* and *hbl-1*. Both single mutants have a rather weak heterochronic phenotype, with only a partial L/A switch happening at the L3-to-L4 molt. At this molt, mutants of *lin-41* have precocious alae in only about 50 % of the animals (Slack et al., 2000) (Figure 3C), and the seam cells of *hbl-1* mutants, although precociously secreting alae, cannot exit the cell cycle and divide again (Abrahante et al., 2003; Lin et al., 2003) (Figure 3C). By contrast, a depletion of both LIN41 and HBL-1 leads to fully penetrant precocious phenotypes at the L3-to-L4 molt, such as complete alae structures on all animals and a complete exit from the cell cycle for all seam cells (Abrahante et al., 2003; Lin et al., 2003) (Figure 3C). Moreover, more than half of these worms display precocious alae even two stages too early (Abrahante et al., 2003; Lin et al., 2003). As these double depletion experiments were performed with reduction-of-function alleles and/or by RNAi, it is unclear what a double null mutant phenotype would look like. It is possible that a full double depletion would lead to a similar phenotype as in *lin-14;lin-28* mutants, with a complete L/A switch happening two stages too early. Because many seam cells of animals depleted for LIN41 and HBL-1 do not go through the symmetric cell division of the L2 stage (Abrahante et al., 2003), the seam cell pattern in Figure 3C is drawn to skip the L2 stage, but this might be oversimplified. The redundancy between *hbl-1* and *lin-41* is emphasized by epistasis experiments, in which double depletions are able to fully suppress both *lin-4* and *let-7* phenotypes, but neither single depletion is able to suppress any *lin-4* phenotypes or fully suppress the *let-7* mutant phenotypes (Abrahante et al., 2003). In summary, the redundancies of the two precocious heterochronic gene pairs *lin-14-lin-28* and *hbl-1-lin-41* have remained unsolved riddles.

1.5 Motivation and aims for this thesis

The heterochronic pathway is partially, if not extensively conserved from *C. elegans* to mammals. This is true for both the involved genes and for their role in controlling stem or progenitor cell divisions, particularly for regulating the switch between a proliferative self-renewal program and differentiation. Next to the conserved miRNAs *let-7* and *lin-4*, the most striking conservation is found for the heterochronic genes encoding the RBPs LIN28 and LIN41. Conserved are not only their domain architecture, but also their function in controlling proliferation versus differentiation programs and their regulation by miRNAs (Ecsedi and Grosshans, 2013; Rehfeld et al., 2015; Shyh-Chang and Daley, 2013). Whereas LIN28 has been extensively studied, much less is known about LIN41, which has only recently been found to promote somatic cell reprogramming to iPSCs (Worringer et al., 2014). Strikingly, the highly conserved *let-7* miRNA regulates LIN41 protein production in different species including worms and mammals. Because *let-7* is known to inhibit self-renewal and promote differentiation in different contexts and organisms, the *let-7*-LIN41 regulatory module has been proposed to be an ancient control mechanism in stem cell development (Ecsedi and Grosshans, 2013).

Despite the interesting conservation of the *let-7*-LIN41 module, little is known about its downstream effectors. In this work, we therefore aimed at identifying the physiologically relevant targets of *C. elegans* LIN41 and the mechanisms through which these are regulated. A first motivation was to find explanations for the crucial role of LIN41 in stem cell proliferation, by identifying potentially conserved key targets of LIN41. The identification of these targets would be of great interest to stem cell research, as it would increase the knowledge about developmental switches from self-renewal to differentiation programs, and as it might help to exploit the therapeutic potential of stem cells in regenerative medicine. A second motivation was to better define the role of *lin-41* in the heterochronic pathway, because it has been unknown how and through which direct targets it controls seam cell proliferation and differentiation.

After identifying interesting LIN41 targets, we defined additional aims for this thesis. First, as our experiments revealed that LIN41 targets only a few mRNAs, a next aim of this work was to characterize those targets, i.e. to define the impact of their regulation through LIN41 on *C. elegans* development. Second, the unexpected finding that LIN41 targets only a specific isoform of LIN-29 hinted towards a potential explanation for the redundancy of *lin-41* with *hbl-1* in the heterochronic pathway. Therefore, we aimed at understanding this redundancy, at the same time hoping to find explanations for additional unsolved riddles of the heterochronic pathway.

2 Results

2.1 Establishing and optimizing ribosome profiling for *Caenorhabditis elegans*

At the heart of this PhD thesis are experiments to globally measure gene expression changes at the level of mRNA translation. Our interest in the miRNA *let-7* and the RNA-binding protein LIN41, shown or proposed to regulate targets at the translational level in *C. elegans* (Ding and Grosshans, 2009; Slack et al., 2000), prompted us to establish the ribosome profiling technique in our lab and for *C. elegans*. In the publication of this section, we present the optimized protocol, established in a joint effort with Andreas Arnold from the lab of Rafal Ciosk. During this effort, we have compared different strategies to purify single ribosomes (monosomes) and different library preparation protocols to clone and sequence ribosome-protected fragments (RPFs). The results of these comparisons, also published in this study, will be helpful to researchers when choosing the optimal strategy depending on the experimental setup. Independent of the chosen monosome purification and library preparation strategy, we further present strategies to optimize the RNase digests and to reduce the ribosomal RNA contaminations in the final sequencing libraries. Together with the presented overview of different ribosome profiling applications, we hope this publication will help researchers inside and outside the *C. elegans* field to perform and optimize the ribosome profiling technique according to their needs.

Publication

Transcriptome-wide measurement of ribosomal occupancy

by ribosome profiling

Methods

Volume 85, September 1 (2015), Pages 75–89

Inferring Gene Regulatory Interactions from Quantitative High-Throughput Measurements

doi link: http://dx.doi.org/10.1016/j.ymeth.2015.06.013

Keywords:

- Ribosome profiling
- Translation
- Translational control
- Post-transcriptional regulation
- Polysome profiling
- *Caenorhabditis elegans*

Highlights:

- Detailed protocol for ribosome profiling experiments in model organisms.
- Presentation of bioinformatic strategies to evaluate ribosome profiling data.
- Comparison of different methods for isolation of single ribosomes.
- Comparison of two ribosome-protected fragment (RPF) library preparation protocols.
- Strategies to improve RPF generation and deplete rRNA fragments.

Methods 85 (2015) 75–89

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Transcriptome-wide measurement of ribosomal occupancy by ribosome profiling

Florian Aeschimann [a], Jieyi Xiong [b], Andreas Arnold [a], Christoph Dieterich [b,*], Helge Großhans [a,*]

[a] Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland
[b] Max Planck Institute for Biology of Ageing, Joseph-Stelzmann-Straße 9b, 50931 Cologne, Germany

ARTICLE INFO

Article history:
Received 16 February 2015
Received in revised form 10 June 2015
Accepted 15 June 2015
Available online 20 June 2015

Keywords:
Ribosome profiling
Translation
Translational control
Post-transcriptional regulation
Polysome profiling
Caenorhabditis elegans

ABSTRACT

Gene expression profiling provides a tool to analyze the internal states of cells or organisms, and their responses to perturbations. While global measurements of mRNA levels have thus been widely used for many years, it is only through the recent development of the ribosome profiling technique that an analogous examination of global mRNA translation programs has become possible. Ribosome profiling reveals which RNAs are being translated to what extent and where the translated open reading frames are located. In addition, different modes of translation regulation can be distinguished and characterized. Here, we present an optimized, step-by-step protocol for ribosome profiling. Although established in *Caenorhabditis elegans*, our protocol and optimization approaches should be equally usable for other model organisms or cell culture with little adaptation. Next to providing a protocol, we compare two different methods for isolation of single ribosomes and two different library preparations, and describe strategies to optimize the RNase digest and to reduce ribosomal RNA contamination in the libraries. Moreover, we discuss bioinformatic strategies to evaluate the quality of the data and explain how the data can be analyzed for different applications. In sum, this article seeks to facilitate the understanding, execution, and optimization of ribosome profiling experiments.

1. Introduction

Microarrays and RNA sequencing methods are powerful tools to profile gene expression at the level of transcript abundance. However, mRNAs can also be regulated at the level of translation, leading to higher or lower amounts of protein than estimated from the mRNA level. A more accurate measure of gene expression thus requires knowledge of a gene's output at the protein level. In principle, this can be obtained by mass spectrometry. However, despite rapid advances [1], mass spectrometry technology does not currently allow the measurement of all the proteins of a cell. By contrast, the recent development of the ribosome profiling technique [2] has enabled global determination of the translational activity of RNAs to the depth and precision of RNA sequencing experiments. This approach can thus improve our understanding of the gene expression status of a cell and reveal specifically which genes are post-transcriptionally regulated at the level of translation, providing a basis for studying the molecular mechanisms of translational control in living cells. Here, in addition to providing a detailed protocol for ribosome profiling and its optimization, we seek to give an overview of some of the scientific questions that can be addressed by this technique, as well as some alternative approaches. It should be emphasized that ribosome profiling is intrinsically limited to measurements of translation rates and thus cannot be used to determine the actual protein levels in a cell, since these are the result of both protein production and turnover rate.

The number of ribosomes that cover a cellular mRNA has long been considered a proxy of its translational efficiency. Before the introduction of ribosome profiling, ribosome coverage was usually analyzed through so-called polysome profiling experiments. After treatment with cycloheximide, a chemical that arrests ribosomes during translation elongation [3,4], lysates are fractionated by ultracentrifugation in linear sucrose density gradients. These gradients separate mRNAs bound by multiple ribosomes (polysomes) in heavier fractions from mRNAs bound by fewer or single ribosomes (monosomes) in lighter fractions. RNA isolation of the different fractions allows the quantification of mRNAs associated with different ribosomal populations by reverse transcription quantitative polymerase chain reaction (RT-qPCR) or Northern blotting. Translational efficiency can then be approximated by calculating the percentage of the mRNA associated with polysomes or

* Corresponding authors.
E-mail addresses: christoph.dieterich@age.mpg.de (C. Dieterich), helge.grosshans@fmi.ch (H. Großhans).

http://dx.doi.org/10.1016/j.ymeth.2015.06.013
1046-2023/

by estimating the ribosome occupancy of the mRNA, i.e. how many ribosomes are bound to this mRNA species on average. By subjecting the RNA from different fractions of such a sucrose gradient to microarrays or RNA sequencing, a more global picture of translational activity can be obtained [5–9]. Nevertheless, with polysome profiling experiments, ribosome occupancy can be quantified with only limited accuracy, because good separation among mRNAs bound to distinct numbers of ribosomes by sucrose gradient centrifugation is only possible for mRNAs associated with few ($\leqslant$5–6) ribosomes. Additionally, high density RNP complexes, termed pseudo-polysomes in one instance [10], can contaminate polysomal fractions. Since polysome profiling experiments also do not provide any information on the position of the ribosomes on the mRNA, they cannot distinguish between ribosomes translating the main open reading frame (ORF) and those located at upstream open reading frames (uORFs). This is problematic, because ribosomes initiating at uORFs are not translating the gene product that is quantified and often even prevent ribosomes from translating the main ORF [11].

A more accurate and global quantification of each gene's average ribosome occupancy can be achieved using the ribosome profiling technique. Since its first description in 2009 [2], this method has been used in a multitude of experimental systems and model organisms [12]. In a ribosome profiling experiment, the mRNA regions bound by ribosomes at the time point of analysis can be sequenced, mapped and quantified (Fig. 1). After treatment with cycloheximide, lysates are incubated with an unspecific RNA endonuclease. Whereas this will cause degradation of 'naked', unstructured mRNA, cycloheximide-arrested ribosomes will physically protect the small mRNA fragments to which they are bound [13]. These surviving mRNA fragments of 28–30 nucleotides are termed ribosome-protected fragments (RPFs). As polysomes will be separated into individual monosomes during the digest, each RPF will occur in a single ribosome. These monosomes can be purified from the rest of the digested lysate by taking advantage of either their high density (sucrose gradients, sucrose cushions) or their large size (size-exclusion chromatography). After RNA extraction from the purified monosomes, the RPFs are separated from other RNAs by size using gel purification, ligated to adapters, reverse transcribed, amplified, and sequenced. In parallel to the RPFs, an aliquot of RNA from the undigested lysate is sequenced to determine mRNA abundance levels. Division of the normalized RPF counts by the normalized RNA sequencing counts mapping to a specific mRNA allows the calculation of ribosome occupancy levels. Similarly to polysome profiling experiments, the ribosome occupancy of an mRNA serves as a proxy for its translational efficiency. In order to use ribosome occupancies for comparing translational efficiencies of RNAs, the speed of translation has to be similar among ribosomes on these RNAs. This criterion is often met when the same transcript is examined under different biological or experimental conditions, so that ribosome occupancy will be a good measure of differential translation of a given mRNA. However, some experimental conditions such as cold shock [14] can change the speed of elongation. Therefore, although it is usually presumed that the speed of elongation is the same between two conditions under test, it should be remembered that, unless tested, this is an assumption rather than a fact. Moreover, the situation is less clear when ribosome occupancy is used to compare translational efficiency among different transcripts, because different features on specific mRNAs can affect translation elongation speed as we will discuss below. Nonetheless, translation rates seem to be consistent between different groups of genes [15], suggesting that this method has some general utility even for inter-transcript comparisons.

Next to providing a measure for translational efficiency, ribosome profiling experiments exhibit positional information about every detected RNA-associated ribosome, i.e., they reveal the specific position of a ribosome on an mRNA. This information can be used to localize ORFs and uORFs with single-nucleotide resolution. Due to this high resolution, it is possible to determine which reading frame of a gene is translated and if there are alternative or overlapping reading frames within one gene [12,15,16]. As an example, one study identified hundreds of transcripts in human monocytes that contain two distinct and active start codons, yielding proteins that may differ in activity due to extended, truncated or out-of-frame ORFs [17]. Ribosome profiling data can also be used to find ORFs that have not been annotated so far. Indeed, many novel ORFs and uORFs have been discovered with ribosome profiling experiments, with the unexpected finding that ribosomes at uORFs often initiate translation at non-AUG initiation codons [2,15,18–21]. In addition, ribosome profiling has enabled the identification of numerous ribosome-associated small ORFs in zebrafish and mammalian cell lines in long noncoding RNAs (lncRNAs) previously thought to have no coding potential [15,22–27]. However, whether ribosomal engagement on these lncRNAs leads to production of functional 'micropeptides' or small proteins, is purely regulatory, or has no function at all has remained a controversial issue. Interestingly, a recent study identified a conserved 46 amino acid micropeptide translated from an annotated lncRNAs that regulates skeletal muscle performance [28]. Additional support for the existence of functional micropeptides stems from studies on uORFs. For instance, in *Arabidopsis thaliana*, the control of expression of downstream main ORFs was shown to depend on the peptide sequences encoded in uORFs [29]. The encoded nascent micropeptides act in *cis* and were proposed to stall ribosomes when interacting with components of the ribosome exit tunnel.

Ribosome profiling data can also provide insights into mechanisms of translational regulation. The number of RPF reads corresponding to a footprint on a specific codon reflects the average time ribosomes spend at this codon. This interpretation can be used to study the speed of the ribosome during the translation process. For example, ribosomes are slowed down at the fifth codon of the reading frame [30], at less frequently occurring codons [31–33] (although see also [34–36] for different conclusions on this issue), or at codons dependent on wobble base-pairing to interact with anti-codons on tRNAs [37]. Ribosomes were also found to slow down when incorporating positively charged amino acids ([32,34], but disputed by [38]), proline-rich sequences [38,39] or other specific amino acid sequences [39] into the nascent polypeptide. Very high peaks in RPF densities at a given codon can be caused by stalled ribosomes during translation elongation or translation termination [15,35,40]. Moreover, in the absence of the protein Dom34, stalled ribosomes can be detected near polyA tails after entering 3′ UTRs [41].

Besides slow or stalled ribosomes, additional aspects of ribosome behavior are revealed by tracing ribosome occupancy along a certain reading frame. For example, ribosomes that change their reading frame during translation elongation or that read through stop codons can be observed in ribosome profiling data [16,42]. RPF densities can also be used to distinguish between several modes of translational inhibition. While an inhibition of translation initiation leads to a uniform decrease of RPFs over the whole ORF, inhibition of translation elongation with an unchanged initiation rate leads to an accumulation of ribosomes at the 5′ end and to a depletion of ribosomes at the 3′ end of the ORF. Similarly, ribosome drop-off reduces the footprints at the 3′ end of the ORF to a larger extent than at the 5′ end of the ORF. Taking advantage of such interpretations, miRNAs in zebrafish and mammalian cells were shown to inhibit translation at the initiation step [43,44], and ribosomes were shown to pause during translation elongation under different stress conditions in yeast and mammalian cells [45–47].

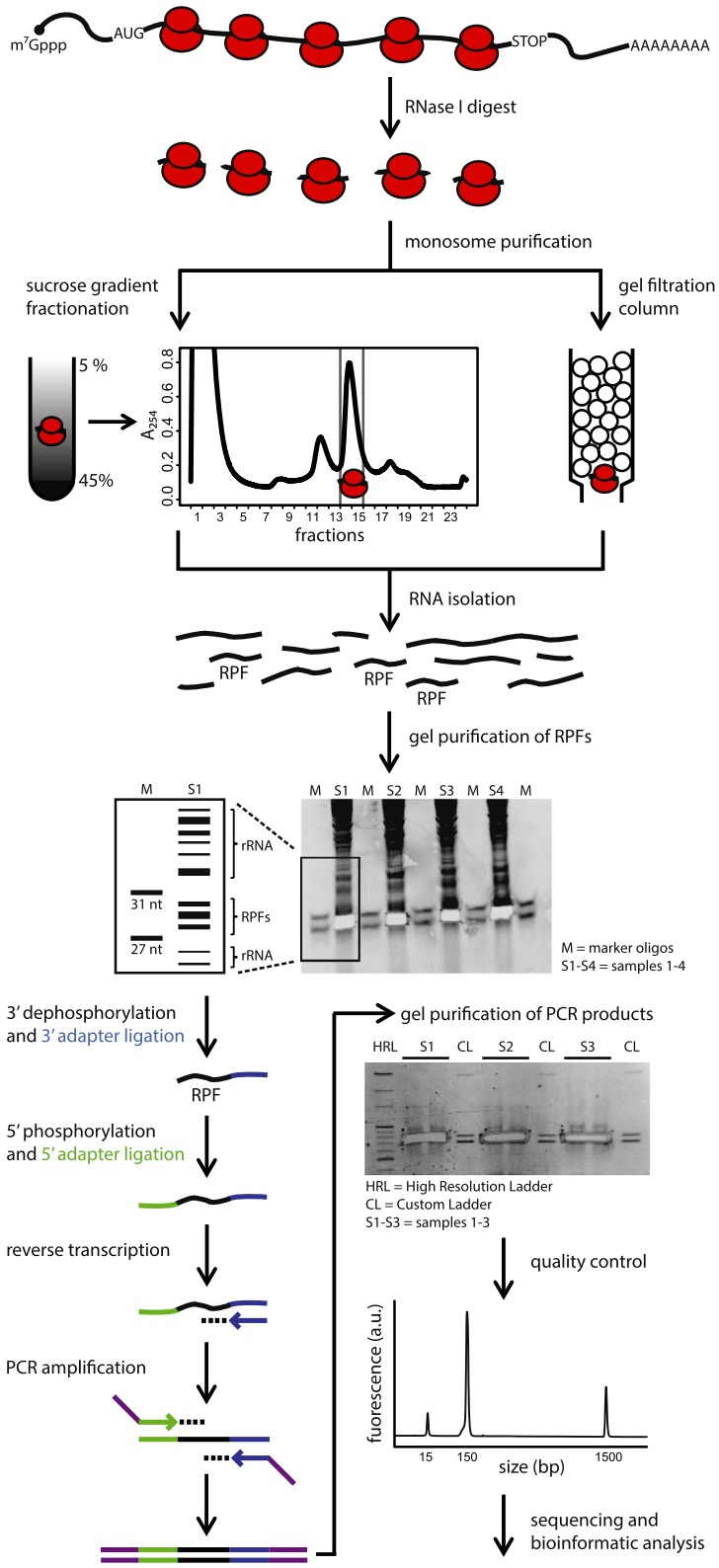


Fig. 1. Outline of a ribosome profiling protocol. Cell or whole-animal lysates are treated with an optimized amount of RNase I to degrade all unprotected RNA. This includes segments of mRNA not bound by ribosomes, thus converting polysomes into an assembly of individual ribosomes (monosomes), each bound to a stable short mRNA fragment termed ribosome-protected fragment (RPF). Monosomes are purified from the rest of the lysate either by fractionation on a linear sucrose gradient or by size-exclusion chromatography using gel filtration columns. RNA is isolated from the purified monosomes and separated by size on a polyacrylamide gel. Marker oligonucleotides of 27 and 31 nucleotides serve to identify a region of the gel that contains the 28–30 nucleotide-long RPFs, which is excised. This step excludes most contaminating ribosomal RNA (rRNA) fragments, which differ in length. To prepare a library for sequencing, RPFs are ligated to a different adapter on each end, reverse transcribed, and amplified. Following gel electrophoresis and extraction of the PCR product, an aliquot of the library is analyzed on an Agilent BioAnalyzer, where a specific peak in the profile at ~150 bp confirms good quality. Following high-throughput sequencing of the library, RPFs are mapped, analyzed, and quantified. In addition to RPFs, an aliquot of RNA from the undigested lysate is sequenced to determine mRNA transcript levels. Small changes in translation efficiency (measured as ribosome occupancy) in different conditions can be reliably and quantitatively detected by averaging all RPF reads of one mRNA and comparing those to its RNA sequencing reads.

Although ribosome profiling experiments monitor the translational activity of individual ribosomes at single-nucleotide resolution, they cannot fully replace polysome profiling experiments for several reasons. First, while ribosome profiling experiments accurately detect changes in ribosomal occupancy of a subset of mRNAs, they are not suited to detect global changes in cellular translation due to the internal normalization of the sequencing reads. This is in contrast to polysome profiles, which directly visualize global changes in translational activity. These profiles can even reveal mechanistic details of such changes, for example by detecting defects in ribosome biogenesis (e.g., Ref. [48]). A second shortcoming of the ribosome profiling technique is that it often cannot distinguish between differentially translated mRNA species originating from the same gene. If transcript variants with different translational activities only differ in their untranslated regions (5′ UTRs and/or 3′ UTRs), RPFs mapping to their coding regions cannot be assigned to a specific variant. In addition, even if only a single mRNA sequence variant exists, it may occur in more than one cellular pool, with subpopulations differing in their translational efficiency, for instance due to differences in subcellular localization. Ribosome profiling can only determine the average number of RPFs per mRNA at the time point of analysis. Thus, it cannot discriminate between, for example, a situation with 10% of a gene's transcripts being highly translated and densely bound by ribosomes while the other 90% of transcripts are ribosome-less, and a situation where 100% of the transcripts are translated at low levels and homogeneously bound by few ribosomes. Therefore, detection of a direct correlation between the number of ribosomes on a message and a feature of interest requires polysome profiling (e.g., Ref. [49]), which unlike ribosome profiling can separate mRNA subpopulations associated with different numbers of ribosomes. Of note, a third possibility to globally measure translated RNAs is affinity purification using tagged-ribosomes, which is discussed and compared to polysome profiling and ribosome profiling in a recent review by King and Gerber [50].

In summary, for mechanistic studies of translational control, ribosome profiling and polysome profiling analyses complement each other. While ribosome profiling can define the positions of individual ribosomes at single-nucleotide resolution, polysome profiling can estimate the translational efficiency of individual transcript subpopulations. For gene expression analysis, ribosome profiling is our method of choice, because it is better suited for high-throughput experiments and more accurately quantifies the average ribosome occupancy of mRNAs.

2. Protocol

The experiment comprises preparation of the biological samples, sample lysis, RNA digest, monosome purification, RNA isolation followed by gel purification of RPFs, and finally library preparation. (The bioinformatic analysis of the data will be covered in Section 3.) After harvesting the biological samples, the whole protocol can be performed in 5 days (see Supplementary Fig. S1 for a flow diagram). However, it is possible to interrupt the protocol at many steps, as indicated in the individual sections. We always perform the lysis and the RNase I digest (step of Section 2.2) plus the monosome purification (step of Section 2.3) on the first day. If purifying the monosomes by sucrose gradient centrifugation (step of Section 2.3.1), day 1 can become a long working day, so we recommend preparing as many things as possible (e.g. equipment, buffers, labeled microfuge tubes) in advance. At the end of day 1, the samples can be stored at −80 °C in Tri Reagent or in isopropanol (precipitation step during the RNA isolation protocol with Tri Reagent). On day 2, we purify the RNA and isolate the RPFs by gel extraction (step of Section 2.4), eluting them from the gel overnight. On day 3, we precipitate the RNA and start the library preparation (step of Section 2.5). Depending on the number of samples and the duration of RNA precipitations (which can always be done overnight) in between the different library preparation steps, we usually finish the library preparation by day 5–7. We normally finish the protocol by performing the gel extraction of the PCR product (step of Section 2.5.6) plus quality control measurements on a separate day, because the gel extraction is the most time consuming step during library preparation. The library preparation for sequencing of the input RNA will take a few additional days, depending on the chosen protocol. Before starting the first experiment, make sure to consult the material safety data sheet of every reagent used in this protocol as well as local health, safety, and environment guidelines, and wear appropriate personal protective equipment. Take special care when using cycloheximide, formamide, Tri Reagent, chloroform, polyacrylamide gels, SYBR-Gold, and SDS.

2.1. Harvest of biological material

To monitor separation of the lysate on a sucrose gradient, it is important to have enough input material. For synchronized *Caenorhabditis elegans* (*C. elegans*) populations, we collect around 50,000 young adult worms, 100,000 larval stage 4 (L4) worms or 200,000 L3 stage worms. We flash-freeze intact worms and add the cycloheximide to arrest the ribosomes only during lysis by cryogenic pulverization (Section 2.2). Flash-freezing followed by pulverization is also a robust and effective cell lysis technique for other model organisms, cultured cells or tissue samples [51]. For any biological material other than *C. elegans* samples, skip steps 1–3 below, but make sure to harvest your biological material in a buffer similar to buffer A described in step 3. For harvesting cultured cells, we refer to the protocol of Ingolia and colleagues [51].

(1) Harvest the worms with M9 buffer (42 mM Na_2HPO_4, 22 mM KH_2PO_4, 86 mM NaCl, 1 mM $MgSO_4$) and transfer them to a 15 ml Falcon tube.
(2) Pellet the worms by centrifugation (1600*g* for 1 min at room temperature) and aspirate off the supernatant. To remove most bacteria, wash the pellet three additional times by addition of M9 buffer, centrifugation and aspiration of the supernatant.
(3) To store the worm pellet in a buffer similar to the assay buffer for ribosome profiling, wash the pellet once with buffer A (20 mM Tris–HCl (pH 7.9 at room temperature), 140 mM KCl, 1.5 mM $MgCl_2$, 0.5% (v/v) Nonidet P40). Transfer the pellet to a microfuge tube, spin down the worms (2000*g* for 30 s at room temperature) and remove as much of the supernatant as possible. Flash-freeze the pellet in liquid nitrogen for storage at −80 °C. In our experience, the frozen worm pellets can be stored at −80 °C for at least three months without any visible decrease of polysomes on the sucrose gradients.

2.2. Lysis and RNase I digest

Due to the instability of DTT (dithiothreitol) and cycloheximide, always freshly prepare the assay buffer (buffer A with 2% (w/v) PTE (polyoxyethylene-10-tridecylether), 1% (w/v) DOC (sodiumdeoxycholate monohydrate), 1 mM DTT, 0.1 mM cycloheximide) on ice with fresh aliquots of DTT and cycloheximide. The concentration of cycloheximide used in this protocol was optimized in polysome profiling experiments of worm lysates and differs from the more commonly used concentration of 0.1 mg/ml (around 0.35 mM). For most applications, both concentrations can similarly be used, but readers should be aware that the concentration of

cycloheximide can have an impact on the distribution of RPFs along the reading frame [52]. To digest the RNA in the lysate, we use the unspecific endoribonuclease RNase I (100 Units/μl, Life Technologies; AM2295). As described in Section 4, it is important to optimize the conditions of the RNase I digest for every experimental system. When using a different unspecific RNase, the library preparation protocol needs to be adjusted. Note that the choice of RNase influences the ribosome profiling data, as exemplified by the enrichment of RPFs in short genes when using RNase I or the enrichment of 3′ UTR reads when using micrococcal nuclease [53]. We do not inhibit RNase I during the isolation of monosomes, as the subsequent addition of Tri Reagent to the purified monosomes rapidly inactivates RNase activity. If for any reason the RNase has to be inactivated prior to the monosome purification, we recommend treatment with an appropriate RNase inhibitor (e.g. SUPERase• In RNase Inhibitor (Life Technologies; AM2696)) rather than with heparin sulfate, because the latter has been shown to inhibit reverse transcription [54,55].

(1) Thaw the frozen pellet on ice for 5–10 min, then add one volume of assay buffer to one volume of pellet (or minimally 300 μl of assay buffer for small pellets). Pre-cool a mortar with liquid nitrogen (fill up completely, let evaporate). Add the pellet to the pre-chilled mortar and crush it with a pestle to fine powder in the presence of liquid nitrogen.
(2) Cover the mortar with aluminum foil to avoid accumulation of condensing water. Let the lysate thaw, either more quickly at room temperature or more slowly on ice if processing many samples in parallel.
(3) Collect the lysate in a microfuge tube and centrifuge at 12,000*g* for 10 min at 4 °C to sediment organelles and cell debris.
(4) Transfer the soluble supernatant to a fresh microfuge tube on ice and measure its absorbance at 260 nm (A_{260}). We usually measure three independent 1:100 dilutions of each lysate and calculate the average A_{260} (using a 1:100 dilution of assay buffer as a blank). To have a rough measure of RNA concentration in the lysate, we calculate the A_{260} Units per ml concentration in the lysate. One A_{260} Unit is defined as the amount of lysate diluted in 1 ml that yields in an absorbance reading of 1.0 with a 10 mm light path length. We therefore calculate the A_{260} Units per ml of lysate as follows:

$$A_{260}\ \text{Units per ml} = (A_{260}(1:100\ \text{dilution}) * 100\ (\text{dilution factor}))/1\ \text{ml}$$

In lysates that we obtained from worm pellets as described in Section 2.1, the A_{260} concentration ranged from 30 to 90 A_{260} Units per ml of undigested lysate.
(5) On ice, dilute the lysate with assay buffer to the preferred concentration (A_{260}/ml) and volume, then add the optimized amount of RNase I (see Section 4). For *C. elegans* lysates, we mix 11 A_{260} Units of lysate with 200 Units of RNase I in a total volume of 385 μl. Incubate the lysate on a heat block for 1 h at 23 °C, 300 rpm. After the incubation, immediately put the samples back on ice and proceed with the monosome purification.
(6) To obtain a sample for isolation of input RNA, add 1 ml of Tri Reagent (Molecular Research Center; TR 118) to the remainder of the undigested lysate. Snap-freeze the sample in liquid nitrogen and store it at −80 °C until RNA extraction.
(7) Perform RNA isolation of the input RNA sample stored in Tri Reagent according to the manufacturer's recommendations. Resuspend the RNA in 50 μl of RNase-free water and store at −80 °C until further use. Isolate rRNA-depleted total RNA or poly-A selected RNA with the method of choice and perform library preparation and high-throughput sequencing according to the manufacturer's protocol of the chosen kits. We usually obtain rRNA-depleted total RNA by performing a 15-min on-column DNase-treatment step using the RNase-Free DNase Set (Qiagen; 79254) and the RNeasy Mini Kit (Qiagen; 74104), followed by treatment with the Ribo-Zero™ rRNA Removal Kit (Epicentre; MRZH11124). For the subsequent library preparation, we use the ScriptSeq v2 RNA-Seq library preparation kit (Epicentre; SSV21124).

2.3. Monosome purification

We commonly use one of two different methods to purify monosomes, linear sucrose density gradients (Section 2.3.1) and size-exclusion chromatography using gel filtration columns for table top centrifuges (Section 2.3.2). Higher quality libraries are obtained with sucrose gradients (see Section 6), but this technique is more time consuming and challenging. The gel filtration columns are easy to use and helpful for processing many samples in parallel. A third option to isolate monosomes are sucrose cushions described in several studies [44,51] and in the manual of the ARTseq™ Ribosome Profiling Kit (Epicentre, RPHMR12126).

2.3.1. Isolation of monosomes by sucrose gradient centrifugation

2.3.1.1. Mixing and ultracentrifugation of the sucrose gradients. This protocol was optimized for ultracentrifugation using a SW 40 rotor and an Optima™L-80 XP Ultracentrifuge (Beckman Coulter). For mixing linear sucrose gradients, we use the Biocomp Gradient Master with Gradient Forming Accessories (Biocomp; 105–914B), i.e. a tube holder, tube caps, and a marker block to designate the level of the heavy-light sucrose interface on the centrifugation tubes.

(1) Prepare filter-sterilized 5% (w/v) and 45% (w/v) sucrose solutions containing 20 mM Tris pH 7.9 (at room temperature), 140 mM KCl and 1.5 mM $MgCl_2$. These sucrose stock solutions can be stored at room temperature for months and supplemented with DTT and cycloheximide just prior to usage. In addition, prepare a filter-sterilized 70% (w/v) sucrose solution for the fractionation step.
(2) To the required amount of 5% (w/v) and 45% (w/v) sucrose solutions (15 ml of each sucrose solution for two gradients), add fresh DTT and cycloheximide to a final concentration of 1 mM and 0.1 mM, respectively. Keep the solutions at room temperature.
(3) Mark polypropylene tubes (14 × 95 mm; Beckman Coulter; 331374) at half their height using the lower edge of the marker block and fill them with 5% (w/v) sucrose solution to just slightly above the mark. To avoid air bubbles, pipette the solution to the wall of the tubes. Fill a syringe with 45% (w/v) sucrose solution and connect it to the cannula for layering provided with the Gradient Master. Insert the cannula quickly to the bottom of the tube and begin layering the heavy solution underneath the light solution. Carefully add the 45% (w/v) sucrose solution until the heavy-light interface rises precisely to the mark in order to have tubes of identical weight for ultracentrifugation. Slowly add the cap of the tube while making sure all the air and the excess sucrose can escape through the cap's overflow valve. Remove the excess 5% (w/v) solution from the central reservoir of the cap. If the gradient contains air bubbles, discard it and start a new gradient.
(4) Mix the gradients at room temperature with the Gradient Master (Biocomp) after selecting the appropriate program (Long Sucr 05–45% wv 1St) and the rotor type (SW 40). Store the gradients at 4 °C until the RNase I digest is finished. Make sure to pre-cool the ultracentrifuge at this point.

(5) After the incubation of the lysate with RNase I, remove the caps from the tubes and load the lysate on top of the gradient. In order to not disturb the gradient, carefully layer the lysate on top of the sucrose solution by slowly pipetting it to the wall of the tube. For *C. elegans* lysates, we load 10 A_{260} Units of lysate in a volume of 350 µl. Place the gradient tubes into pre-chilled rotor buckets and centrifuge at 4 °C for 3 h at 39,000 rpm. After the centrifugation, carefully remove the buckets from the rotor and store them at 4 °C or on ice, taking care not to disturb the gradients. Proceed to the fractionation as soon as possible.

2.3.1.2. Fractionation of the sucrose gradients. We fractionate the gradients by pumping with 70% (w/v) sucrose using a Tris peristaltic Pump (Teledyn ISCO), a Gradient Fractionator (Brandel; BR-184-X) and a fraction collector (Gilson; FC-203B). To record absorbance profiles at 254 nm, we utilize an Econo UV monitor EM-1 (Biorad) coupled to a LabJack U6 data acquisition device using the DAQFactory-Express software. The absorbance at 254 nm monitors the presence of RNA across the gradient and enables identification of the fractions containing the monosomes. To get a continuous profile, we set the sampling rate of the data acquisition to 10 A_{254} measurements per second. Before starting the first fractionation, check the pump speed and the settings of the fraction collector. The pump speed, selected in percentage of maximal speed for the Tris Pump, should be adjusted to get a flow rate of around 1 ml in 30 s using the 70% (w/v) sucrose solution to push. Once set to a certain percentage, the pump speed usually does not change unless one of the tubings is exchanged. In order to obtain clean monosome fractions, the 12 ml gradients are collected in 24 fractions of equal volume (0.5 ml per fraction). With a flow rate of 1 ml in 30 s, the fraction collector should be set to 15 s per tube.

(1) Switch on the UV detector about 30 min before starting the fractionation and, if using the Econo UV monitor, set the sensitivity to 0.5. Set the baseline of the UV monitor by running water through the system. Remove the water again from all tubings by pushing air through the system. Pump the 70% (w/v) sucrose solution through the tubing until the first drops come out of the needle. This will prevent air bubbles from disturbing the gradient during the fractionation.
(2) Fill the fraction collector with open microfuge tubes. Attach a gradient tube to the gradient fractionator system, fixing it in an upright position (grease the borders of the centrifugation tube to prevent leaking). With the needle, pierce the tube from the bottom and raise the needle until its opening is entirely within the tube. Start the data acquisition and unload the gradient by pumping 70% (w/v) sucrose. Start the fraction collector as soon as the first drop reaches it. When the last fraction is collected, stop the pump, the fraction collector and the data acquisition. Store the collected fractions on ice. If desired, the fractions can also be snap-frozen in liquid nitrogen and stored at −80 °C overnight. Pump in reverse direction to re-collect the 70% (w/v) sucrose solution and clean each tubing with water before filling it with air again. Repeat this step for every gradient. In our experience, this procedure takes around 15–20 min per gradient.
(3) Analyze the acquired data and pool the two fractions corresponding to the monosome peak (normally fractions 13–14 or fractions 14–15, as depicted in Fig. 1). Split the mixed sample into two microfuge tubes (0.5 ml each) and add 1 ml of Tri Reagent to each tube. Flash-freeze the samples in liquid nitrogen and store them at −80 °C or proceed with the RNA isolation. Keep one of the tubes as a back-up at −80 °C in case of problems during the RPF isolation or library preparation.
(4) To clean all material of the sucrose, place all the parts used for the gradient preparation in a beaker full of deionized water over night. Clean each tubing with sterile 1% (w/v) SDS and store it in 30% (v/v) ethanol to prevent bacterial growth.

2.3.2. Isolation of monosomes by size-exclusion chromatography

To purify monosomes with gel filtration columns for table-top centrifuges, we use MicroSpin S-400 HR columns (GE Healthcare, 27-5140-0) and slightly adapt the protocol described in the manual of the ARTseq Ribosome Profiling Kit (Epicentre). Be sure to wash the columns properly before applying the digested lysate, since remaining storage buffer decreases the yield of RPFs, likely due to the disassembly of ribosomes through EDTA.

(1) Resuspend the resin of the columns by brief vortexing. Open the column lids by turning at least one quarter-turn and twist off the bottom closure. Place the column in a collection tube and spin for 1 min at 600*g*, discard the flow-through. (Note that columns should be drained by gravity flow according to the Epicentre manual; however, this failed to work in our hands.)
(2) Wash the column 3 times by adding 500 µl buffer A to the column, resuspending the resin by brief vortexing, centrifuging 1 min at 600*g* and discarding the flow-through. In the last wash, centrifuge for 4 min at 600*g*.
(3) Place the column into a fresh 1.5 ml microfuge tube. Apply 100 µl of the RNase-digested lysate to the top-center of the resin bed and spin for 2 min at 600*g* to collect the monosomes. Add 1 ml of Tri Reagent to the flow-through and flash-freeze the samples in liquid nitrogen to store them at −80 °C, or proceed with the RNA isolation. As a backup, steps 1–3 can be repeated with a new gel filtration column and another 100 µl of the same lysate.

2.4. Isolation of the RPFs

For extraction of the 28–30 nucleotides long RPFs, we separate the RNA from the isolated monosomes using Novex 15% (w/v) Polyacrylamide TBE-Urea Gels (Life Technologies; EC6885BOX). As size markers for extraction of the RPFs, we recommend use of a 27 and a 31 nucleotide-long RNA oligonucleotide. For higher cutting accuracy, we load these size markers in each neighboring well of every sample well (Fig. 1). From this step of the protocol until cDNA is obtained in the step of Section 2.5.5, precautions have to be taken to protect the RNA from RNase contaminations. All materials should either be certified RNase-free or treated with an RNase decontamination solution such as RNaseZAP (Life Technologies; AM9780). The reagents and buffers should contain only RNase-free components, including the water. We use commercially available RNase-free water (Life Technologies; AM99397) when using small amounts and treat deionized water with diethylpyrocarbonate (DEPC) for large amounts [56]. In addition, protect the samples from human RNases by wearing clean gloves at all times.

(1) Perform RNA isolation of the monosome fraction samples stored in Tri Reagent according to the manufacturer's recommendations. Resuspend the RNA in 15 µl of RNase-free water. Proceed to the gel extraction or store the RNA at −80 °C until further use.

(2) To keep the electrophoresis apparatus free of RNase contamination, clean it with RNaseZAP (Life Technologies; AM9780). Assemble the apparatus, add running buffer prepared with DEPC-treated water and wash out the remaining storage buffer from all the wells of the gel. Prerun the gel at 170 V for 10–15 min.
(3) Add 15 μl of formamide loading buffer (formamide with 0.05% (w/v) bromophenol blue) to each 15 μl RNA sample. Prepare a marker mix with 1 μl of each RNA marker oligo (100 ng/μl) mixed with 8 μl RNase-free water and 10 μl of formamide loading buffer per well. Denature RNA (samples and marker mix) for 90 s at 80 °C and put back on ice. Load the RNA samples (30 μl) on the gel, with 20 μl of the marker mix loaded in each neighboring well. With this setup, a maximum of 4 different RNA samples can be loaded per gel.
(4) Run the gel at 170 V for 1.5 h (or until the bromophenol blue band is close to running out of the gel). To visualize the RNA, incubate the gel in 1× SYBR-Gold (Life Technologies; S-11494) diluted in running buffer (in DEPC-treated water) for about 8 min at room temperature. Transfer the gel to a transparent RNase-free surface (e.g. plastic wrap or glass plate).
(5) With a sterile scalpel, excise the 28–30 nucleotide region of each sample lane under UV light (this corresponds to the region between the two RNA marker oligos). Transfer each excised gel piece to a Gel Breaker tube (IST Engineering; 3388-100) inside an RNase-free 2 ml microfuge tube.
(6) Centrifuge for 2 min at full speed to force the gel through the holes, fragmenting it into small pieces. Invert the Gel Breaker tube to add any remaining larger gel pieces to the lower 2 ml microfuge tube. Add 600 μl of freshly prepared RNase-free cracking buffer (20 mM Tris–HCl (pH 7.9), 1 mM EDTA, 400 mM Ammoniumacetate, 0.5% (w/v) SDS) to the gel debris and snap-freeze the samples in liquid nitrogen. When totally frozen, remove the tubes from the liquid nitrogen and rotate them gently at room temperature overnight to elute the RNA.
(7) Spin the eluate through a 5 μm filter tube (IST Engineering; 5388-50) at full speed and room temperature for 2 min. If using the provided 1.5 ml microfuge tube to collect the filtrate, we recommend transferring the eluate to a different 1.5 ml RNase-free microfuge tube for RNA precipitation (pellets do not form nicely in the provided tubes).
(8) Add 2 μl of GlycoBlue (15 mg/ml; Life Technologies; AM9515) and 750 μl of isopropanol to the filtrate, mix well and incubate on dry ice for at least 30 min or at −80 °C overnight to precipitate the RNA.
(9) Pellet the precipitated RNA by centrifugation at 12,000*g*, 30 min, 4 °C. The RNA pellets are very small at this step and may even be invisible if dispersed.
(10) Remove the supernatant, add 950 μl of 75% (v/v) ethanol and invert the tube to wash the pellet. (Do not resuspend to keep an intact pellet!) Spin down the pellet for 5 min at 7500*g*, 4 °C. Carefully pipette away all ethanol, air-dry the pellet for about 5 min (take care not to over-dry the pellet) and resuspend in 43 μl of RNase-free water. Proceed with the next step or store the RNA at −80 °C.

2.5. Library preparation for ribosome-protected fragments

We have tried two different strategies for library preparation of RPFs. Although both strategies gave similar results, we prefer the strategy described below (see Section 6). We base our library preparation on the TruSeq Small RNA Sample Preparation Kit (Illumina; RS-200-0012), but have to modify the 5′ and 3′ ends of the RPFs in order to allow ligation of the adapters. RNase I produces RNA fragments with 5′ hydroxyl ends and 3′ ends with a 2′, 3′ cyclic phosphodiester. Both ends can be modified by treatment with T4 polynucleotide kinase (NEB; M0201S). The 3′ hydroxyl group required for the 3′ adapter ligation step is obtained using its phosphatase activity (in the absence of ATP), while the 5′ phosphate required for the 5′ adapter ligation is obtained using its kinase activity (in the presence of ATP).

2.5.1. 3′ dephosphorylation

(1) Denature RNA (43 μl) at 80 °C for 90 s and transfer back on ice. To dephosphorylate the RNA, add 5 μl of T4 PNK buffer (10×, NEB), 1 μl of T4 polynucleotide kinase (10 U/μl; NEB; M0201S) and 1 μl of RNasin Ribonuclease Inhibitor (40 U/μl; Promega; N2111). Incubate the reaction for 1 h at 37 °C.
(2) To precipitate the RNA, add 1 μl of GlycoBlue (15 mg/ml; Life Technologies; AM9515), 309 μl of RNase-free water and 40 μl of 3 M sodium acetate (Life Technologies; AM9740) before mixing the reactions with 500 μl of isopropanol. Incubate on dry ice for at least 30 min or at −80 °C overnight. Recover the RNA as described in steps 9–10 of Section 2.4, but resuspend it in 5 μl of RNase-free water. We often get dispersed pellets after isopropanol precipitation at this step. Therefore, make sure to wash all the RNA off the wall of the tube.

2.5.2. Ligation of 3′ adapter

This step is adapted from the Illumina® TruSeq™ Small RNA Sample Preparation protocol. Unless specified otherwise, the reagents are included in the TruSeq Small RNA Sample Preparation Kit (Illumina; RS-200-0012).

(1) Mix the RNA (5 μl) with 1 μl of RNA 3′ Adapter (RA3), denature at 80 °C for 90 s and transfer back on ice.
(2) Prepare the ligation master mix with 2 μl Ligation Buffer (HML), 1 μl RNase inhibitor (40 U/μl) and 1 μl T4 RNA Ligase 2, truncated (200 U/μl; NEB; M0242S) per sample, mix by pipetting. For several samples, add 10% overage to the master mix.
(3) Add 4 μl of the master mix to each sample (6 μl), mix by pipetting and incubate for 1 h at 28 °C. After 1 h of incubation, add 1 μl of Stop Solution (STP) to each reaction and gently mix by pipetting. Incubate another 15 min at 28 °C, then place the tubes on ice.
(4) To precipitate the RNA, add 1 μl of GlycoBlue (15 mg/ml; Life Technologies; AM9515), 348 μl of RNase-free water and 40 μl of 3 M sodium acetate (Life Technologies; AM9740) before mixing the reactions with 500 μl of isopropanol. Incubate on dry ice for at least 30 min or at −80 °C overnight. Recover the RNA as described in steps 9–10 of Section 2.4, but resuspend it in 37.5 μl of RNase-free water.

2.5.3. 5′ phosphorylation

(1) Denature RNA (37.5 μl) at 80 °C for 90 s and transfer back to ice. To phosphorylate the RNA, add 5 μl of T4 PNK buffer (10×, NEB), 5 μl of 10 mM ATP, 1.5 μl of T4 polynucleotide kinase (10 U/μl; NEB; M0201S) and 1 μl of RNasin Ribonuclease Inhibitor (40 U/μl; Promega; N2111). Incubate the reaction for 30 min at 37 °C before heat-inactivating the enzyme for 10 min at 70 °C.
(2) To precipitate the RNA, add 1 μl of GlycoBlue (15 mg/ml; Life Technologies; AM9515), 309 μl of RNase-free water and 40 μl of 3 M sodium acetate (Life Technologies; AM9740)

before mixing the reactions with 500 µl of isopropanol. Incubate on dry ice for at least 30 min or at −80 °C overnight. Recover the RNA as described in steps 9–10 of Section 2.4, but resuspend it in 7 µl of RNase-free water.

2.5.4. Ligation of 5′ adapter

This step is adapted from the Illumina® TruSeq™ Small RNA Sample Preparation protocol. All reagents are included in the TruSeq Small RNA Sample Preparation Kit (Illumina; RS-200-0012). We add the Stop Solution (STP) to the ligation mix for historical reasons and have never tested if it is needed for efficient 5′ adapter ligation. In the TruSeq Small RNA Sample Preparation Kit, STP is added to stop the 3′ adapter ligation and therefore included in the 5′ adapter ligation mix.

(1) Separately denature the RNA (7 µl) and the required volume (1.1 µl per sample) of RNA 5′ Adapter (RA5) at 80 °C for 90 s and transfer back to ice.
(2) Prepare the ligation master mix with 2 µl Ligation Buffer (HML), 1 µl Stop Solution (STP), 1 µl RNA 5′ Adapter (RA5), 1 µl ATP (10 mM), 1 µl RNase inhibitor (40 U/µl) and 1 µl T4 RNA Ligase per sample, mix by pipetting. For several samples, add 10% overage to the master mix.
(3) Add 7 µl of the master mix to each sample (7 µl), mix by pipetting and incubate for 1 h at 28 °C, then place the tubes on ice. Store the remaining RNA not used for the reverse transcription at −80 °C as a back-up.

2.5.5. Reverse transcription and PCR amplification

This step is adapted from the Illumina® TruSeq™ Small RNA Sample Preparation protocol. Unless specified otherwise, the reagents are included in the TruSeq Small RNA Sample Preparation Kit (Illumina; RS-200-0012). The number of PCR cycles was optimized for our *C. elegans* samples. For any other input material, we recommend to optimize the number of PCR cycles as described in the protocol of Ingolia and colleagues [51].

(1) Mix 6 µl of ligated RNA with 1 µl of the RNA RT Primer (RTP), denature in a nuclease-free PCR tube at 80 °C for 90 s and transfer back to ice.
(2) Prepare the reverse transcription master mix with 2 µl 5× First Strand Buffer, 0.5 µl 12.5 mM dNTP mix (1:1 dilution of the supplied 25 mM mix), 1 µl 100 mM DTT, 1 µl RNase inhibitor (40 U/µl) and 1 µl SuperScript II Reverse Transcriptase (200 U/µl; Life Technologies; 18064-014) per sample, mix by pipetting. For several samples, add 10% overage to the master mix.
(3) Add 5.5 µl of the master mix to each sample (7 µl), mix by pipetting and incubate for 1 h at 50 °C, then place the tubes on ice. Store the remaining cDNA not used for PCR amplification at −20 °C as a back-up.
(4) To 2.5 µl of each cDNA, add 18.5 µl deionized water, 25 µl PCR Mix (PML), 2 µl RNA PCR Primer (RP1) and 2 µl RNA PCR Primer Index (RPIX), mix by pipetting.
(5) Perform the PCR amplification using the following program: (a) 30 s at 98 °C, (b) 9 cycles of: 10 s at 98 °C, 30 s at 60 °C, 15 s at 72 °C, (c) 10 min at 72 °C, (d) hold at 4 °C. Store the PCR reactions at −20 °C or proceed to gel extraction of the PCR product.

2.5.6. Gel extraction of the PCR product

This step is adapted from the Illumina® TruSeq™ Small RNA Sample Preparation protocol. Unless specified otherwise, the reagents are included in the TruSeq Small RNA Sample Preparation Kit (Illumina; RS-200-0012). For gel extraction of the PCR products, we use Novex 6% (w/v) Polyacrylamide TBE Gels (Life Technologies; EC6865BOX).

(1) Add 10 µl of 6× LDye (Loading dye) to the 50 µl PCR products. Per gel, mix 1 µl High Resolution Ladder (HRL) with 1 µl 6× LDye. Per sample loaded, mix 1 µl Custom Ladder (CL) with 1 µl 6× LDye.
(2) Prepare the gel according to the manufacturer's recommendations. Load 2 µl of the High Resolution Ladder (HRL) in the first well, always load 2 times 30 µl of each PCR product in two adjacent wells and separate different PCR products with a well loaded with 2 µl of the Custom Ladder (CL). Per gel, a maximum of 3 different PCR reactions can be loaded. Run the gel for 50 min at 145 V.
(3) Stain the DNA by incubation of the gel in 1× SYBR-Gold (Life Technologies; S-11494) diluted in running buffer for about 3 min at room temperature. Place the gel on a clean glass or plastic surface.
(4) With a sterile scalpel, excise the two adjacent bands of each sample running between the 145 bp and 160 bp bands of the Custom Ladder (CL) under UV light (Fig. 1). Transfer the gel pieces to Gel Breaker tubes (IST Engineering; 3388-100) inside 2 ml microfuge tubes.
(5) Centrifuge for 2 min at full speed to force the gel through the holes, fragmenting it into small pieces. Add 300 µl of deionized water to the gel debris and rotate the tubes gently at room temperature for at least 2 h to elute the DNA.
(6) Spin the eluate through a 5-µm filter tube (IST Engineering; 5388-50) at 600*g* and room temperature for at least 10 s. To precipitate the DNA, add the filtrate, 2 µl of GlycoBlue (15 mg/ml; Life Technologies; AM9515), 30 µl of 3 M sodium acetate (Life Technologies; AM9740), and 975 µl of pre-chilled (−20 °C) 100% ethanol to a fresh 1.5 ml microfuge tube. Immediately pellet the precipitated DNA by centrifugation at 20,000*g*, 20 min, 4 °C.
(7) Remove the supernatant and add 950 µl of 75% (v/v) ethanol to wash the pellet. Spin down the pellet for 5 min at 7500*g*, 4 °C. Carefully pipette away all ethanol, air-dry the pellet and resuspend it in 15 µl of deionized water.
(8) Check the quality of the library by running 1 µl of each library DNA on a DNA1000 Chip on the Agilent 2100 Bioanalyzer according to the manufacturer's protocol. The profile of a successfully prepared library shows a single peak that normally runs at ∼150 bp (Fig. 1). Proceed to high-throughput sequencing or store the libraries at −20 °C.
(9) The libraries are now ready for quantification (e.g. with a Qubit Fluorometer), equimolar pooling (for multi-plexing), clustering and high-throughput sequencing with the Illumina system according to the manufacturer's protocols. A read length of 50 bp is sufficient for libraries with ribosome-protected fragments. For *C. elegans* samples, we usually aim at 45–50 million clusters per sample and run 4 libraries per lane on an Illumina HiSeq2000 machine.

3. Bioinformatic analysis and quality control

The general analysis pipeline for ribosome-protected fragments is based on RNA-Seq analysis (Fig. 2A). As a first step, sequencing adaptors should be removed from the 3′ end of all short reads with dedicated software such as Flexbar [57]. Unwanted contaminants such as rRNA reads should be removed before the actual mapping step. To this end, we employ conventional short read mappers such as Bowtie2 [58], BWA [59] or SOAP2 [60] with custom-build index structures. Typically, these pre-processed reads will subsequently be mapped to a reference genome by using splice-aware mapping software like Tophat

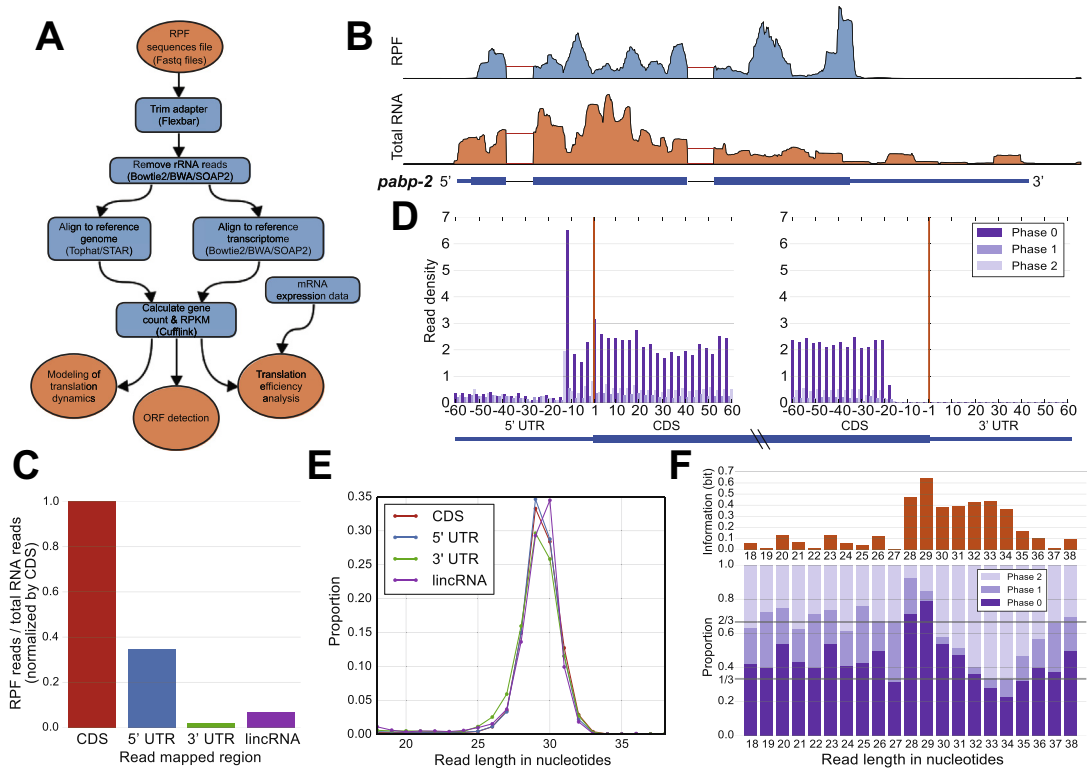


Fig. 2. RPF read features. (A) Flow chart of RPF data analysis (see main text for details). (B–F) Analysis of merged reads from ten ribosome profiling and total RNA sequencing libraries, respectively, prepared according to the protocol described in this manuscript. The libraries were obtained from synchronized wild-type *C. elegans* populations harvested at ten different time points over 18 h of development (GSE52905). (B) Exemplary RPF read coverage (blue profile) and total-RNA read coverage (orange profile) of the *C. elegans* gene *pabp-2*. Red lines represent junction reads aligned by splice-aware mapping software. The annotation track is plotted at the bottom. Thick blue lines/medium blue lines/thin black lines represent CDS/UTRs/introns, respectively. (C) Relative RPF read densities in different gene regions. Included are reads uniquely mapped in CDS/UTR regions of 20,611 coding genes (as annotated in Wormbase WS236), or in 161 lincRNAs annotated in Ensembl. The height of bars is the relative ratio of RPF read number vs. total RNA read number in each region. All the ratios are normalized to the CDS ratio (left most bar), which has been set to 1. (D) Metagene analysis of the read 5′ end coverage around the translation initiation site (left; 3,751 loci) and the translation stop site (right; 11,006 loci). Selected were all loci with 5′ UTRs or 3′ UTRs longer than 60 nucleotides for translation initiation and translation stop sites, respectively. Only 5′ ends of uniquely mappable 29-nucleotide RPF reads are shown. The orange vertical lines show the translation initiation and the translation stop sites in the left and right figures, respectively. (E) Length distributions of RPF reads uniquely mapped in CDS/UTR regions and lincRNAs, as in panel C. (F) Periodicity of RPF reads with different fragment sizes. Only 5′ ends of non-junction reads that uniquely map to CDS regions are shown. Top: The information content of phase frequency distributions as calculated by $\log_2$ 3-H, where H is the Shannon entropy calculated by $-\sum_{i=0}^{2} f_i \times \log_2 f_i \cdot f_i$ is the relative frequency of reads in phase i. Bottom: Relative frequencies of reads in different coding phases.

[61] and STAR [62]. Alternatively, RPF reads could be mapped to the corresponding transcriptome reference directly with short read mapping software. However, we recommend using the first approach as it encompasses the whole available sequence space. Since RPF reads are relatively short (<50 nucleotides), we recommend to use a stringent mapping parameter set (i.e. allowing no more than one mismatch and considering only uniquely mapping reads). Finally, read counts and RPKM values on sequence features of interest may be obtained with conventional software solutions such as Cufflinks [63] or HT-Seq [64].

3.1. RPF read features

RPF reads have several properties that distinguish them from conventional RNA-seq data with regard to, first, read coverage profile along transcript body, second, read length distribution (i.e. footprint size), and, third, periodicity pattern of read densities in translated regions. All of these features are indicative of good quality RPF data. As mentioned before, RPFs can be utilized not only to quantify average translation, but also to identify positional biases in ribosome occupancy or to detect novel translated regions in gene bodies such as uORFs. Here, we use a *C. elegans* ribosome profiling data set that we previously published [65] to illustrate the utility of the aforementioned RPF read features. The respective data sets are available from the GEO database through accession GSE52905. The 10 *C. elegans* samples constitute a time-series collected at every 2 h during development from L3 to young adult stage. We merged all 10 samples together to maximize read coverage over all expressed transcripts.

3.1.1. RPF read coverage

A typical example of RPF read coverage along a gene body is shown in Fig. 2B for the gene *pabp-2*. A direct comparison of RPF read coverage (blue profile) and total RNA-seq read coverage (orange profile) reveals a specific lack of RPF read coverage from the 3′ UTR. Fig. 2C depicts a global statistics on these coverage differences between coding sequence (CDS), UTR and non-coding RNA sequence features. In our data, most (∼99%) of the mapped RPF reads align to coding regions as previously reported [2]. In particular, the 3′ UTR is devoid of RPFs, consistent with release of the 80S ribosomes at the translation stop site. Guttman and colleagues [24] have proposed that the dramatic coverage drop between CDS and 3′ UTR regions is characteristic of genuine translation and distinguishes protein-coding RNAs from lincRNAs, which may be decorated with RPF reads, but do not show this positional drop in coverage. They captured this difference in a Ribosome Release Score (RRS) to distinguish between the two classes of genes. However, it is important to note that the number of RPFs detected in 3′ UTRs is dependent on the type of RNase used as well as the conditions of the digest [53].

Different coverage shapes of the CDS region may reflect biological features such as accumulation of ribosomes at the translational start site or at ribosome pause sites during translation elongation or termination [15]. A good example of that can be studied in a metagene analysis of our *C. elegans* data: the positional bias around the translation start site (Fig. 2D). Herein, we see a distinct peak of 5′-ends of RPF reads at 12 nucleotides upstream of the translation initiation codon. This indicates that the P and A sites of the ribosome are located at nucleotides 13–15 and 16–18, respectively, counted from the 5′-end of 29-nucleotide reads. This has been similarly observed with 28-nucleotide RPF reads in a previous yeast study [2].

3.1.2. Characteristic read length distribution

The mRNA fragments that are protected by ribosome complexes have a specific length distribution. RPF are typically enriched in the 28–30 nucleotide size range [2] or 20–22 nucleotide size range [66] depending on the usage of different elongation inhibitors. Moreover, different buffer conditions such as high or low ion strength may alter the characteristic size distribution [51]. Here we only discuss ribosome profiling with reads that fall into the 28–30 nucleotide size range, which are relevant for the most commonly used experimental method, i.e. blocking with cycloheximide. In our *C. elegans* samples, ~97% of all mapped reads fall into the size range of 26–32 nucleotides (Fig. 2E). Previously, characteristic read length distributions have been used to define a fragment length organization similarity score (FLOSS) [23]. This score reflects differences in the 26–33 nucleotide RPF read length distribution of true ribosome protected vs. other transcript regions. This score was then used together with ORF predictions to identify novel translated regions even outside of annotated genes.

3.1.3. Periodicity in RPF read coverage

The eukaryotic 80S ribosome shifts along its RNA template with a step size of 3 nucleotides (one codon) in every elongation cycle. Accordingly, a distinctive feature of RPF data from coding regions is a triplet periodicity pattern that reflects this discrete movement [2]. In our *C. elegans* data, we could confirm that reads with different fragment lengths show distinct frequency preferences over the three possible reading frames (Fig. 2F). RPF reads with length between 28 and 34 nucleotides have higher phase information than all other sizes, yet 29-nucleotide RPF reads are the most informative as expressed by their deviation from a uniform distribution across all three phases (see Fig. 2F). This periodic pattern of 29-nucleotide fragments has been exploited to identify novel coding regions [22].

3.2. Further analysis of RPF data

RPF data can be used not only to study the dynamic process of translation or to discover novel peptide-coding sequences, but also to estimate translation efficiency [15,44]. Translation elongation rates are remarkably constant over different gene classes and independent of length and protein abundance [15]. Given constant elongation rates, translation efficiency of an ORF can be readily estimated by dividing the normalized RPF read counts by the normalized mRNA read counts. However, this basic estimate is necessarily correlated with cytosolic mRNA levels [67]; a more accurate estimation would eliminate this dependency on mRNA levels through the analysis of partial variance [67]. Recently, Olshen et al. developed a statistical model (Babel) to detect significant changes in translation regulation for individual transcripts between two conditions based on an errors-in-variables regression model [68]. Herein, both RPF and RNA count data are modeled by negative binomial distributions and measurement errors are acknowledged in both data types (dependent and independent variable). This strategy is a methodological improvement over simple linear regression, which may suffer from regression dilution (i.e. biasing the regression slope towards smaller absolute values). However, a systematic assessment of performance gains by independent experiments is difficult and has not been published to our knowledge.

4. Optimizing the RNase digest

When first establishing ribosome profiling in an experimental system, it is important to optimize the RNase digest. On the one hand, overdigestion of the RNA in the lysate will disrupt the structure of the ribosomes and therefore could lead to degradation of the mRNA fragments bound by ribosomes. On the other hand, when insufficiently digesting the RNA in the lysate, the ribosome-protected fragments will not be trimmed to their characteristic size and will not all migrate to the region excised from the gel. Therefore, in order to obtain 28–30 nucleotide-long RPFs for each piece of RNA bound by a ribosome, the concentration of the RNase has to be optimized. To find the optimal quantity of the RNA endonuclease RNase I in our experimental system, a fixed amount of *C. elegans* lysate was incubated for 1 h at 23 °C with increasing RNase I concentration and analyzed by sucrose gradient ultracentrifugation. In the undigested lysate, a large fraction of the ribosomes are bound to mRNAs harboring several ribosomes. These polysomes migrate at higher density on the sucrose gradient than mRNAs bound to one (monosomes) or two (disomes) ribosomes (Fig. 3A). A digest with a low concentration of RNase I (125 Units) separated most of the ribosomes from each other, resulting in a clear decrease of polysomes and a clear increase of monosomes in the lysate (Fig. 3B). Using 250 U of RNase I led to a similar profile compared to the 125 U digest (Fig. 3C). Higher RNase I amounts, such as 500 U, caused an apparent decrease of monosomes, most likely due to degradation of the ribosomes themselves (Fig. 3D). A digest with even higher amounts of RNase I led to an almost complete degradation of the ribosomes (Fig. 3E and F). We define the optimal RNase concentration as the highest concentration at which the monosome peak is still at its maximum size. With the depicted and other RNase I digestion series, we found an addition of 200–250 U of RNase I to be optimal for a digest of a *C. elegans* lysate of 11 A_{260} Units in a volume of 385 µl. We performed the digests with this amount and dilution of lysate because we load 10 A_{260} Units in a volume of 350 µl onto the sucrose gradient for ribosome profiling. The optimized conditions correspond to an RNase I concentration of 520–650 Units per ml of lysate, with a lysate of 28.6 A_{260} Units per ml. However, we note that we have never determined by library preparation that more extensive digestion does indeed decrease RPF quality.

5. The problem of ribosomal RNA contamination

The main contaminant in RPF libraries is RNase I resistant ribosomal RNA (rRNA), which can easily make up more than 80% of the reads [2], leading to extensive loss of informative RPF reads. Different strategies have been described to decrease the rRNA contamination in order to sequence more RPFs and thereby get higher coverage. A popular strategy is to treat the isolated RNA with an Epicentre Ribo-Zero™ rRNA Removal Kit, either before (e.g., as described in the protocol for the ARTseq™ Ribosome Profiling Kit) or after (e.g., Ref. [22]) purification of the RPFs by polyacrylamide gel electrophoresis (PAGE). Another strategy is to pull-down the most abundant rRNA sequences occurring in the libraries with biotinylated complementary oligonucleotides [51]. However, all these methods eliminate only a part of the rRNA reads

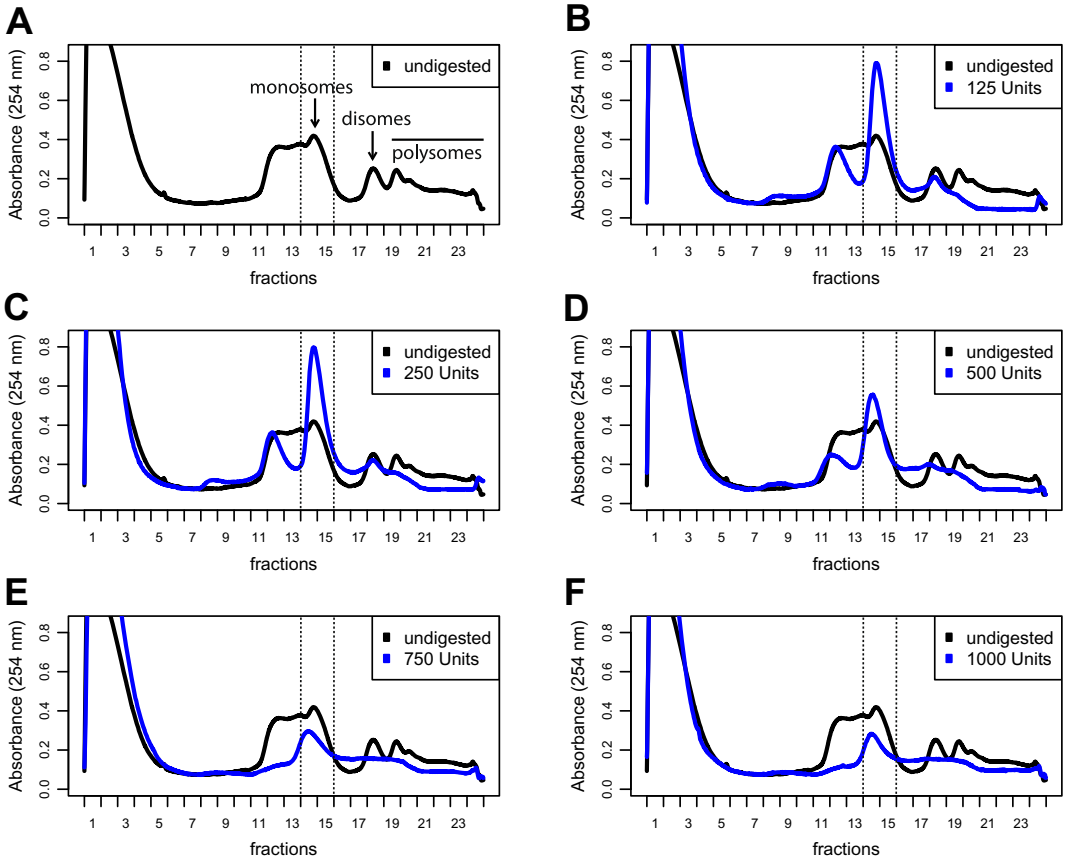


Fig. 3. Optimizing RNA digest with a series of increasing RNase I concentrations. Polysome profiles of linear 5–45% (w/v) sucrose gradients were recorded as described in the protocol. The 12 ml gradients were loaded with 10 A_{260} Units of *C. elegans* larval stage L4 lysates (in 350 µl) after digesting the RNA in the lysates with increasing amounts of RNase I Units. The optimal amount of RNase I under these conditions was found to be 250 U, because this was the highest amount of RNase that yielded a large monosome peak. For ribosome profiling, RNA is extracted from fractions 14 and 15, which contain the monosome peak (indicated with dashed lines).

and many published ribosome profiling experiments were performed without any steps to remove rRNA contaminants. During our efforts to reduce rRNA reads in the *C. elegans* libraries, we found that optimizing the cutting accuracy when PAGE-purifying the RPFs significantly improves their purity. In our initial experiment, we excised the region ranging from 26 to 34 nucleotides, as suggested earlier [51]. The fraction of mRNA reads was only about 15% with most (85%) of the reads originating from ribosomal RNA (Fig. 4A). [The data obtained and analyzed in this and the following Section 6 have been deposited in the GEO database under the series accession number GSE65948.] Interestingly, most of the rRNA reads were longer than 32 nucleotides, whereas the mRNA reads almost exclusively originated from reads 28–30 nucleotides in length. We also noticed that the highest rRNA contamination derived from 35 nucleotide long reads, although most 35 nucleotides long RNAs were removed by cutting the 26–34 nucleotide window from the gel. To compare different cutting boundaries for gel extraction of the RPFs, we divided up the same RNase I digested lysate to excise the following regions: (i) 28–30 nucleotides, (ii) 25–30 nucleotides, (iii) 28–35 nucleotides (Fig. 4B–D). After performing the library preparation and sequencing the libraries, we found that the 28–35 nucleotide region resulted in only 9% of mRNA reads due to massive rRNA contamination from 35 nucleotides long reads (Fig. 4D). By contrast, excluding RPFs above 30 nucleotides in length from the gel greatly increased the fraction of mRNA reads. While the 28–30 nucleotide region resulted in 57% of mRNA reads (Fig. 4C), the 25–30 nucleotide region yielded 62% of mRNA reads (Fig. 4B). These percentages are very similar, because only few of the stable rRNA fragments are shorter than 34 nucleotides, as visible in the sequenced reads (Fig. 4) and on the gel in Fig. 1. Thus, simply by cutting out a narrower region of RNA fragments, we could increase the mRNA fraction in the library about 4-fold compared to excision of the classical 26–34 nucleotide window. Importantly, every experimental system will bear different RNase-resistant rRNA fragments. For example, in yeast, the main rRNA contaminants seem to be 24–25 nucleotides long [2]. Although this experiment was performed with a *C. elegans* lysate with most RNase-resistant rRNA fragments clearly differing in size from the RPFs, optimizing the cutting accuracy to a narrow window will likely improve the quality of any RPF library.

6. Comparison between different methods to isolate monosomes and between different library preparations

To determine our preferred ribosome profiling protocol, we compared the quality of two different methods for isolation of monosomes and two different library preparations. Isolating monosomes by sucrose gradient centrifugation and fractionation is very labor-intensive and limits the number of samples that can be processed in parallel. With our equipment, the procedure takes more than 4 h with a maximum of 6 samples. The ARTseq™ Ribosome Profiling Kit (Epicentre, RPHMR12126) describes the option to purify monosomes using MicroSpin S-400 HR gel filtration columns (GE Healthcare, 27-5140-0). This procedure normally takes less than 1 h and can process up to 24 samples in parallel. To evaluate if we could replace the sucrose gradient fractionation

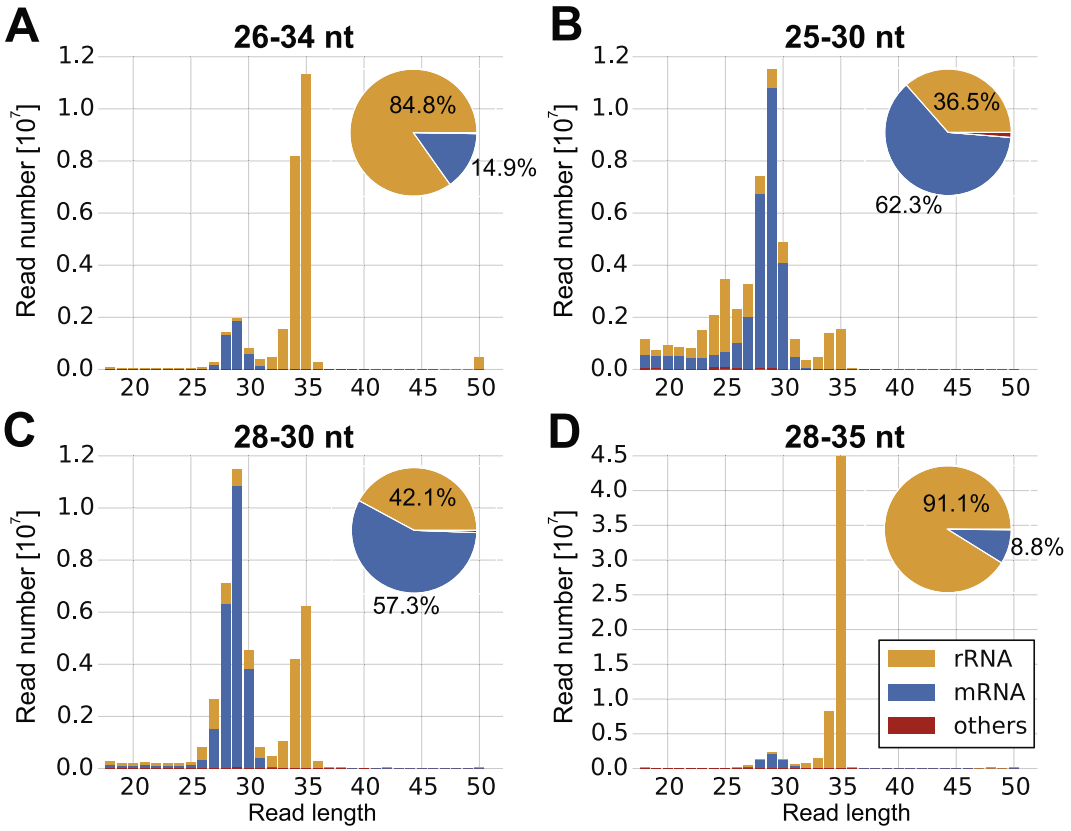


Fig. 4. Optimizing the cutting accuracy for PAGE purification of RPFs. Reads mapping to mRNA, rRNA, or any other RNA (others) in each library as percentage of total (pie charts) and as a function of the read length (histograms). (A) The originally described [2,51] gel purification of the 26–34 nucleotide (nt) region of the RNA isolated from purified monosomes yielded about 85% of contaminating rRNA reads in the library. (B–D) To test if more accurate cutting can improve separation of RPFs from rRNA contaminants, the same RNA sample from purified monosomes of a larval stage L4 *C. elegans* lysate was divided up to excise three different regions from the gel. While including 35 nucleotides long RNAs increased the rRNA contamination to over 90%, omitting the region above 30 nucleotides greatly decreased the ribosomal RNA contamination in the library.

(labeled "SG" in Fig. 5) by size-exclusion chromatography using gel filtration columns (labeled "GC" in Fig. 5), we compared the two monosome isolation methods with the same biological sample. In addition, we compared the RPF library preparations of our protocol (labeled "LP1" in Fig. 5) and of the ARTseq™ Ribosome Profiling Kit (Epicentre, RPHMR12126, labeled "LP2" in Fig. 5), omitting the step with the Ribo-Zero™ Magnetic Kit for rRNA removal for better comparison. In our opinion, both library preparations are very similar in terms of workload and hands-on time. Whereas the ARTseq™ Ribosome Profiling Kit has only one RNA ligation step and therefore no 5′ phosphorylation step, it additionally contains a PAGE purification step for the cDNA and a circularization step. In the experiment illustrated in Fig. 5, we performed both monosome isolation methods twice from the same RNase I digested lysate (see steps of Section 2.3.1 and 2.3.2 of the protocol), prepared from young adult *C. elegans* worms. We then pooled each pair of RNA samples isolated with the same monosome purification method. To compare the two different library preparations in parallel with the same input RNA from each monosome isolation method, we again split each pooled sample into two to obtain four different libraries. All libraries resulted in similarly strong bands after PCR amplification and there was no difference in quality of the gel-purified library DNA on the Agilent 2100 Bioanalyzer (data not shown). After mapping the sequenced reads, we found that the quality of the RPFs was good in all four libraries, as judged by the percentage of mRNA reads mapping to coding regions (>99.5% in all libraries) and by the phasing of the 29 nucleotides long reads (Fig. S2A). Next, we counted the read numbers mapping to the coding sequences of each gene and analyzed how well those correlated among the libraries (Fig. 5A). The correlations between the different monosome isolation methods followed by the same library preparation were higher ($r > 0.98$) than the correlations between the different library preparations following the same monosome isolation method ($r > 0.96$). This also holds true when analyzing the correlations of only the 29 nucleotides long RPF reads (Fig. S2B). We conclude that it is important to stick to one library preparation protocol for related experiments in order to more reliably compare the different data sets.

When looking at the reads in more detail, we found that the size distribution of the reads was similar in all libraries, but the percentage of contaminating rRNA and therefore the RPF sequencing depth differed between the libraries (Fig. 5B). Purifying monosomes by sucrose gradient fractionation clearly removed more contaminating rRNA from the samples compared to the purification by gel filtration columns. This is expected because we can separate the specific monosome fraction from higher and lower density fractions in a sucrose gradient, whereas the gel filtration columns remove only low molecular weight complexes from the sample. We also found more rRNA reads in the libraries prepared with the ARTseq™ Ribosome Profiling Kit (LP2) compared to our protocol (LP1). A closer look revealed that this difference mainly comes from the slightly broader range of fragment sizes in the LP2 libraries, resulting in a higher abundance of rRNA reads longer than 32 nucleotides. This broader RPF size range does not reflect cutting accuracy of RPFs from the gel, because the same gel-purified RNA input was subjected to both library preparations. In order to investigate possible causes of elevated abundance of 33–34 nucleotides long rRNA reads in the LP2 libraries, we selected

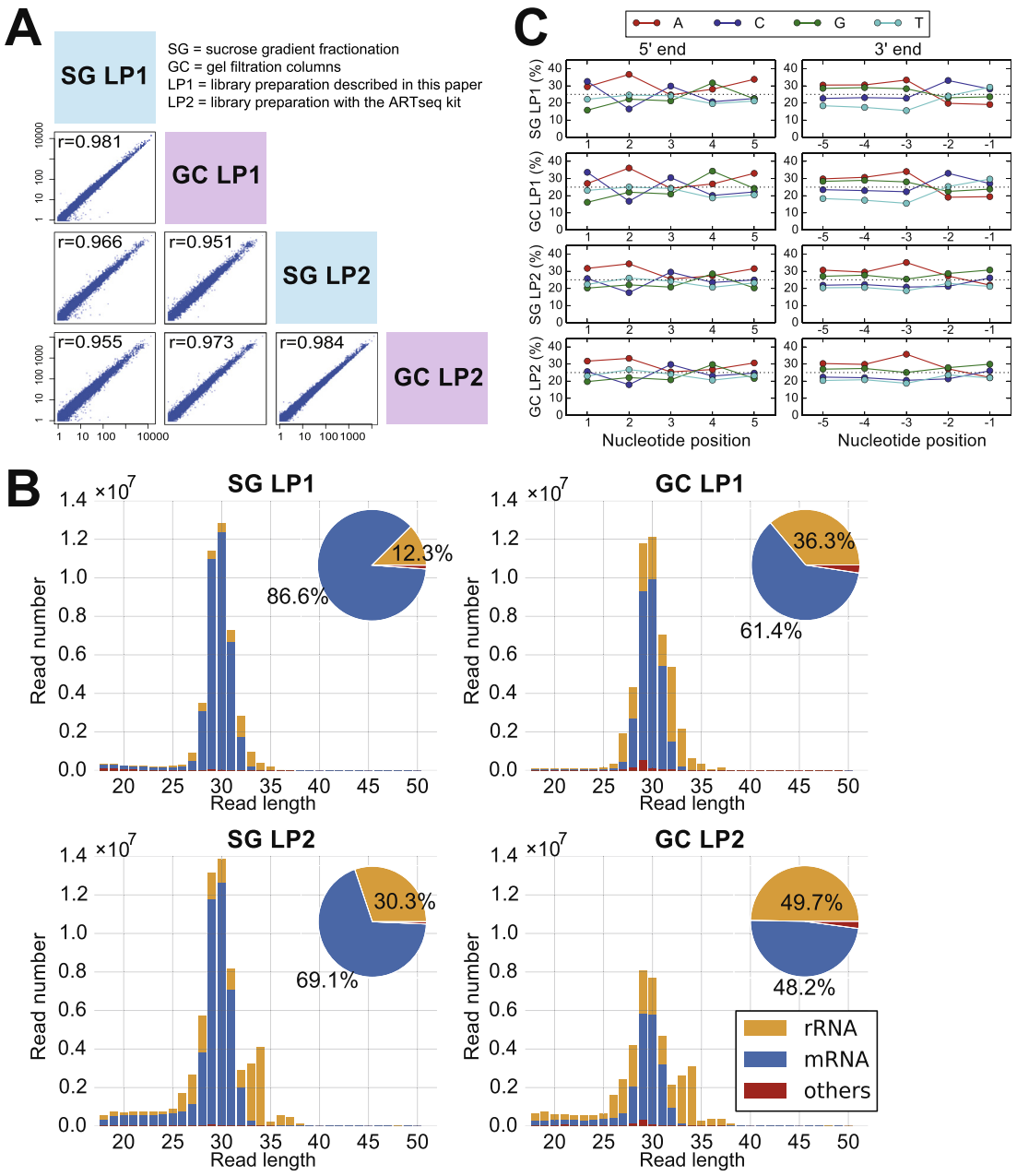


Fig. 5. Comparison between two different methods for isolation of monosomes and two different library preparations. The same *C. elegans* lysate was subjected to two different methods of monosome purification, sucrose gradient fractionation (SG) or size exclusion chromatography using gel filtration columns (GC). RNA from both monosome samples was then prepared for sequencing using the library preparation protocol described in this paper (LP1) or the library preparation protocol of the ARTseq™ Ribosome Profiling Kit (LP2). RPF libraries were prepared from young adult worms, which generally results in a higher percentage of mRNA reads when compared to libraries prepared from larval stage L4 worms (our unpublished observations, also cp. Fig. 4). (A) Pairwise correlation plot for the number of reads mapping to the CDS of every gene on a $\log_{10}$ scale. (B) Reads mapping to mRNA, rRNA or any other RNA (others) in each library as percentage of total (pie charts) and as a function of the read length (histograms). (C) The proportions (in percent) of each nucleotide at the first five positions from the 5′ end and 3′ end of all reads mapping to coding sequence. Position "1" indicates the first position from the 5′ end, whereas position "−1" indicates the first position from the 3′ end.

only the reads of this size and compared the read numbers mapping to different genes in the different protocols. When comparing the LP2 samples to the LP1 samples, we found one very clearly enriched rRNA, gene F31C3.7, with almost all the reads mapping to the same location within the gene (data not shown). We do not know why this particular RNase-resistant rRNA sequence is preferentially enriched in the LP2 samples compared to the LP1 samples. One possibility is a cloning bias, since different library preparations can have different biases for certain RNA sequences or structures, as described below.

Sequence biases in small RNA sequencing libraries have been shown to be predominantly caused by the bias of RNA ligases for certain sequence or structural features [69–71]. The library preparation of the ARTseq™ Ribosome Profiling Kit is based on RNA ligation of a preadenylylated linker followed by reverse transcription and circularization, as optimized and described by Ingolia and colleagues [51]. This strategy was found to reduce sequence biases in the libraries, mainly due to the optimized 3′ linker ligation [2]. However, the intramolecular ligation to circularize the cDNA was found to further reduce sequence biases in the library [2]. Our library preparation protocol also includes ligation of a preadenylylated linker to the 3′ end of RPFs, but differs in the 5′ adapter ligation, as we perform an intermolecular RNA ligation. Therefore, we wanted to compare the ligation biases for the four protocols and

analyzed the RPF nucleotide proportions at nucleotide positions one through five from both the 5′ and the 3′ ends (Figs. 5C and S2C). The ligation bias at the 3′ terminal nucleotide was minor, but differed between the two library preparation protocols. At the 5′ end, the two library preparations resulted in slightly more pronounced nucleotide biases, with the LP1 protocol having a bias for the nucleotide C and the LP2 protocol having a bias for the nucleotide A (Figs. 5C and S2C). However, the biases at the 5′ terminal nucleotide were comparable in extent and not higher than at other positions within the RPFs, suggesting that the ligation bias does not have a major impact on the quality of these libraries. Nevertheless, we also compared the read coverage for some example genes and found that the coverage appeared more even with the library preparation of the LP2 protocol (see Fig. S3 for four example genes). Since we did not observe major nucleotide biases at the 5′ or 3′ terminal positions of the RPFs, it is possible that structural features within the RPFs lead to ligation biases that are reflected in these coverage profiles [69,71].

To examine the homogeneity of read coverage in the different libraries globally, we determined the RPF read coverage (RPKM) at every nucleotide position of the CDS of each mRNA. Next, we calculated the mean and standard deviation for the read coverage per nucleotide position for each CDS. As a measure of coverage homogeneity, we calculated the coefficient of variation (CV) by dividing the standard deviation by the mean. A higher CV reflects a greater dispersion in the distribution of read coverages at nucleotide positions and therefore corresponds to a lower coverage homogeneity. This analysis revealed comparable homogeneity when preparing samples by sucrose gradient and gel filtration (Fig. S4). For highly expressed genes, the CV was lower in the libraries prepared with the LP2 protocol than with the LP1 protocol, confirming that the LP2 protocol leads to a more even read coverage. For genes at low to moderate expression levels (up to $10^{1.4}$ RPKM, Fig. S4), which constitute the bulk of active genes (Table S1), the CV was generally and in all protocols much higher. The negative correlation of CV with expression level is a consequence of the dependency of the CV on the mapped read numbers. Due to this dependency, differences in the CVs for genes at low to moderate expression levels likely do not represent different coverage homogeneities, but can be explained by the sequencing depths of the different libraries (Table S2, number of uniquely mapped CDS reads). The increased coverage homogeneity of the LP2 protocol may be beneficial when performing, for instance, studies on transcript features that may influence the speed of ribosomes. However, it is less important for relative quantification in gene expression studies. For the choice of protocol for gene expression studies, we thus consider the observed differences in percentage of mRNA reads (Fig. 5B) or CDS reads (Table S2) within all mapped reads to be at least as important as the observed difference in coverage homogeneity.

In summary, we obtained good quality RPF libraries with no more than 50% rRNA reads with all four protocols. The addition of an rRNA removal step with a Ribo-Zero™ Magnetic Kit would likely further decrease the rRNA read fraction in all libraries, especially for experiments with species that are supported by the kit. For the highest enrichment of RPFs in the library, we suggest using sucrose gradients for isolation of monosomes. This increases the RPF coverage and therefore allows higher multiplexing to reduce sequencing cost. However, gel filtration columns are much easier to handle, do not need any additional equipment and allow the processing of many samples in parallel in a short time. For labs not equipped with polysome profiling infrastructure or for large-scale experiments in general, we therefore recommend using gel filtration columns to purify monosomes. For subsequent library preparation, both tested protocols performed well. While the protocol described in this article led to lower rRNA contamination in the library, the protocol of the ARTseq™ Ribosome Profiling Kit led to a more even read coverage.

7. Concluding remarks

In this article, we describe our optimized protocol for obtaining ribosome profiling libraries, followed by our bioinformatic strategies to analyze RPF data. Our protocol covers the standard technique, but can be extended with the rapidly developing amendments or specializations that have evolved due to the wide interest in the ribosome profiling technique. Some recent technical innovations promise a higher level of specificity for RPFs through affinity purification of ribosomes. Addition of a ribosome affinity purification step after isolation of monosomes helps to remove background reads from pseudo-RPFs, e.g. originating from RNAs protected by RNA-binding proteins outside the ribosome structure [23]. Similarly, affinity-tagged ribosomes allow ribosome profiling experiments with membrane-bound ribosomes on specific subcellular compartments, such as the endoplasmic reticulum or mitochondria [72,73]. Other subpopulations of translating ribosomes have been assayed by purifying ribosomes associated with specific accessory factors or by immunoprecipitation of tags within the nascent peptide itself [74,75]. Furthermore, the ribosome profiling technique has been diversified through the application of different drugs and by purifying RPFs of different sizes. Cycloheximide is the most widely used drug to arrest translating ribosomes, but ribosome profiling can also be performed without drug treatment or after treatment with other translation elongation inhibitors such as emetine [15,23]. Other drug treatments can provide ribosome profiling data sets with different meanings, as illustrated by the use of harringtonine, lactimidomycin or high doses of puromycin, which lead to an accumulation of ribosomes at translation initiation sites, and thus reveal where 80S ribosomes are assembled on the mRNAs [15,19,20]. In addition, by purifying 20–22 or 16 nucleotides long RPFs, ribosomes can be monitored in a different conformation during the translation elongation cycle or when arrested at the 3′ end of truncated RNAs, respectively [41,66]. With all these recent developments, ribosome profiling experiments have revealed novel insights into mechanisms of translation and future alterations of the technique could help to unravel many mysteries associated with translating ribosomes.

Author contributions

F.A. designed research, performed and analyzed experiments, and wrote the manuscript. J.X. performed all computational analysis and contributed to writing the manuscript. A.A. designed research, and performed and analyzed experiments. C.D. supervised computational analysis and contributed to writing the manuscript. H.G. designed research, analyzed experiments, and wrote the manuscript.

Conflict of interest statement

The authors have no conflicts to declare.

Transparency Document

The Transparency document associated with this article can be found in the online version.

Acknowledgements

We thank Peggy Janich and Manuel de la Mata for comments on the manuscript, Ariel Bazzini, Rafal Ciosk, and Dimos Gaidatzis for discussion and help with protocol development, Kirsten Jacobeit, Sophie Dessus-Babus, Stéphane Thiry, and Tim Roloff for advice on, and help with, library preparation and high-throughput sequencing. H.G. has received funding for the research leading to these results from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 241985 (European Research Council 'miRTurn'), the NCCR RNA & Disease funded by the Swiss National Science Foundation, and the Novartis Research Foundation through the FMI. The funding sources had no role in study design; execution; the writing of the report; nor the decision to submit it for publication.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ymeth.2015.06.013.

References

[1] A.I. Nesvizhskii, Nat. Methods 11 (2014) 1114–1125.
[2] N.T. Ingolia, S. Ghaemmaghami, J.R.S. Newman, J.S. Weissman, Science 324 (2009) 218–223.
[3] T. Schneider-Poetsch, J. Ju, D.E. Eyler, Y. Dang, S. Bhat, W.C. Merrick, R. Green, B. Shen, J.O. Liu, Nat. Chem. Biol. 6 (2010) 209–217.
[4] N. Garreau de Loubresse, I. Prokhorova, W. Holtkamp, M.V. Rodnina, G. Yusupova, M. Yusupov, Nature 513 (2014) 517–522.
[5] Y. Arava, Y. Wang, J.D. Storey, C.L. Liu, P.O. Brown, D. Herschlag, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 3889–3894.
[6] D.G. Hendrickson, D.J. Hogan, H.L. McCullough, J.W. Myers, D. Herschlag, J.E. Ferrell, P.O. Brown, PLoS Biol. 7 (2009) e1000238.
[7] L. Spangenberg, P. Shigunov, A.P. Abud, A.R. Cofre, M.A. Stimamiglio, C. Kuligovski, J. Zych, A.V. Schittini, A.D. Costa, C.K. Rebelatto, P.R. Brofman, S. Goldenberg, A. Correa, H. Naya, B. Dallagiovanna, Stem Cell Res. 11 (2013) 902–912.
[8] F.V. Karginov, G.J. Hannon, Genes Dev. 27 (2013) 1624–1632.
[9] V. Gandin, K. Sikstrom, T. Alain, M. Morita, S. McLaughlan, O. Larsson, I. Topisirovic, J. Vis. Exp. (2014).
[10] R. Thermann, M.W. Hentze, Nature 447 (2007) 875–878.
[11] K. Wethmar, Wiley Interdiscip. Rev. RNA 5 (2014) 765–778.
[12] N.T. Ingolia, Nat. Rev. Genet. 15 (2014) 205–213.
[13] S.L. Wolin, P. Walter, EMBO J. 7 (1988) 3559–3569.
[14] J.R. Knight, A. Bastide, A. Roobol, J. Roobol, T.J. Jackson, W. Utami, D.A. Barrett, C.M. Smales, A.E. Willis, Biochem. J. 465 (2015) 227–238.
[15] N.T. Ingolia, L.F. Lareau, J.S. Weissman, Cell 147 (2011) 789–802.
[16] A.M. Michel, K.R. Choudhury, A.E. Firth, N.T. Ingolia, J.F. Atkins, P.V. Baranov, Genome Res. 22 (2012) 2219–2229.
[17] S.W. Brubaker, A.E. Gauthier, E.W. Mills, N.T. Ingolia, J.C. Kagan, Cell 156 (2014) 800–811.
[18] G.A. Brar, M. Yassour, N. Friedman, A. Regev, N.T. Ingolia, J.S. Weissman, Science 335 (2012) 552–557.
[19] S. Lee, B. Liu, S. Lee, S.X. Huang, B. Shen, S.B. Qian, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) E2424–E2432.
[20] C. Fritsch, A. Herrmann, M. Nothnagel, K. Szafranski, K. Huse, F. Schumann, S. Schreiber, M. Platzer, M. Krawczak, J. Hampe, M. Brosch, Genome Res. 22 (2012) 2208–2218.
[21] N. Stern-Ginossar, B. Weisburd, A. Michalski, V.T. Le, M.Y. Hein, S.X. Huang, M. Ma, B. Shen, S.B. Qian, H. Hengel, M. Mann, N.T. Ingolia, J.S. Weissman, Science 338 (2012) 1088–1093.
[22] A.A. Bazzini, T.G. Johnstone, R. Christiano, S.D. Mackowiak, B. Obermayer, E.S. Fleming, C.E. Vejnar, M.T. Lee, N. Rajewsky, T.C. Walther, A.J. Giraldez, EMBO J. 33 (2014) 981–993.
[23] N.T. Ingolia, G.A. Brar, N. Stern-Ginossar, M.S. Harris, G.J.S. Talhouarne, S.E. Jackson, M.R. Wills, J.S. Weissman, Cell Rep. 8 (2014) 1365–1379.
[24] M. Guttman, P. Russell, N.T. Ingolia, J.S. Weissman, E.S. Lander, Cell 154 (2013) 240–251.
[25] G.L. Chew, A. Pauli, J.L. Rinn, A. Regev, A.F. Schier, E. Valen, Development 140 (2013) 2828–2834.
[26] I. Ulitsky, D.P. Bartel, Cell 154 (2013) 26–46.
[27] S. Quinodoz, M. Guttman, Trends Cell Biol. 24 (2014) 651–663.
[28] D.M. Anderson, K.M. Anderson, C.L. Chang, C.A. Makarewich, B.R. Nelson, J.R. McAnally, P. Kasaragod, J.M. Shelton, J. Liou, R. Bassel-Duby, E.N. Olson, Cell 160 (2015) 595–606.
[29] I. Ebina, M. Takemoto-Tsutsumi, S. Watanabe, H. Koyama, Y. Endo, K. Kimata, T. Igarashi, K. Murakami, R. Kudo, A. Ohsumi, A.L. Noh, H. Takahashi, S. Naito, H. Onouchi, Nucleic Acids Res. 43 (2015) 1562–1576.
[30] Y. Han, X. Gao, B. Liu, J. Wan, X. Zhang, S.B. Qian, Cell Res. 24 (2014) 842–851.
[31] J. Gardin, R. Yeasmin, A. Yurovsky, Y. Cai, S. Skiena, B. Futcher, Elife 3 (2014).
[32] A. Dana, T. Tuller, PLoS Comput. Biol. 8 (2012) e1002755.
[33] S. Pechmann, J.W. Chartron, J. Frydman, Nat. Struct. Mol. Biol. 21 (2014) 1100–1105.
[34] C.A. Charneski, L.D. Hurst, PLoS Biol. 11 (2013) e1001508.
[35] G.W. Li, E. Oh, J.S. Weissman, Nature 484 (2012) 538–541.
[36] W. Qian, J.R. Yang, N.M. Pearson, C. Maclean, J. Zhang, PLoS Genet. 8 (2012) e1002603.
[37] M. Stadler, A. Fire, RNA 17 (2011) 2063–2073.
[38] C.G. Artieri, H.B. Fraser, Genome Res. 24 (2014) 2011–2021.
[39] C.J. Woolstenhulme, S. Parajuli, D.W. Healey, D.P. Valverde, E.N. Petersen, A.L. Starosta, N.R. Guydosh, W.E. Johnson, D.N. Wilson, A.R. Buskirk, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) E878–887.
[40] S. Elgamal, A. Katz, S.J. Hersch, D. Newsom, P. White, W.W. Navarre, M. Ibba, PLoS Genet. 10 (2014) e1004553.
[41] N.R. Guydosh, R. Green, Cell 156 (2014) 950–962.
[42] J.G. Dunn, C.K. Foo, N.G. Belletier, E.R. Gavis, J.S. Weissman, Elife 2 (2013) e01179.
[43] H. Guo, N.T. Ingolia, J.S. Weissman, D.P. Bartel, Nature 466 (2010) 835–840.
[44] A.A. Bazzini, M.T. Lee, A.J. Giraldez, Science 336 (2012) 233–237.
[45] M.V. Gerashchenko, A.V. Lobanov, V.N. Gladyshev, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 17394–17399.
[46] R. Shalgi, J.A. Hurt, I. Krykbaeva, M. Taipale, S. Lindquist, C.B. Burge, Mol. Cell 49 (2013) 439–452.
[47] B. Liu, Y. Han, S.B. Qian, Mol. Cell 49 (2013) 453–463.
[48] A. Fernandez-Pevida, O. Rodriguez-Galan, A. Diaz-Quintana, D. Kressler, J. de la Cruz, J. Biol. Chem. 287 (2012) 38390–38407.
[49] A.O. Subtelny, S.W. Eichhorn, G.R. Chen, H. Sive, D.P. Bartel, Nature 508 (2014) 66–71.
[50] H.A. King, A.P. Gerber, Brief Funct. Genomics (2014).
[51] N.T. Ingolia, G.A. Brar, S. Rouskin, A.M. McGeachy, J.S. Weissman, Nat. Protoc. 7 (2012) 1534–1550.
[52] M.V. Gerashchenko, V.N. Gladyshev, Nucleic Acids Res. 42 (2014) e134.
[53] T.P. Miettinen, M. Bjorklund, Nucleic Acids Res. 43 (2015) 1019–1034.
[54] M.J. del Prete, R. Vernal, H. Dolznig, E.W. Mullner, J.A. Garcia-Sanz, RNA 13 (2007) 414–421.
[55] M.L. Johnson, C. Navanukraw, A.T. Grazul-Bilska, L.P. Reynolds, D.A. Redmer, Biotechniques 35 (2003) 1140–1142. 1144.
[56] M.R. Green, J. Sambrook, Cold Spring Harbor Laboratory Press, New York, (2012) 345–454.
[57] M. Dodt, J.T. Roehr, R. Ahmed, C. Dieterich, Biology (Basel) 1 (2012) 895–905.
[58] B. Langmead, S.L. Salzberg, Nat. Methods 9 (2012) 357–359.
[59] H. Li, R. Durbin, Bioinformatics 25 (2009) 1754–1760.
[60] R. Li, C. Yu, Y. Li, T.W. Lam, S.M. Yiu, K. Kristiansen, J. Wang, Bioinformatics 25 (2009) 1966–1967.
[61] C. Trapnell, L. Pachter, S.L. Salzberg, Bioinformatics 25 (2009) 1105–1111.
[62] A. Dobin, C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T.R. Gingeras, Bioinformatics 29 (2013) 15–21.
[63] C. Trapnell, B.A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M.J. van Baren, S.L. Salzberg, B.J. Wold, L. Pachter, Nat. Biotechnol. 28 (2010) 511–515.
[64] S. Anders, P.T. Pyl, W. Huber, Bioinformatics 31 (2015) 166–169.
[65] G.J. Hendriks, D. Gaidatzis, F. Aeschimann, H. Grosshans, Mol. Cell 53 (2014) 380–392.
[66] L.F. Lareau, D.H. Hite, G.J. Hogan, P.O. Brown, Elife 3 (2014).
[67] O. Larsson, N. Sonenberg, R. Nadon, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 21487–21492.
[68] A.B. Olshen, A.C. Hsieh, C.R. Stumpf, R.A. Olshen, D. Ruggero, B.S. Taylor, Bioinformatics 29 (2013) 2995–3002.
[69] M. Hafner, N. Renwick, M. Brown, A. Mihailovic, D. Holoch, C. Lin, J.T. Pena, J.D. Nusbaum, P. Morozov, J. Ludwig, T. Ojo, S. Luo, G. Schroth, T. Tuschl, RNA 17 (2011) 1697–1712.
[70] A.D. Jayaprakash, O. Jabado, B.D. Brown, R. Sachidanandam, Nucleic Acids Res. 39 (2011) e141.
[71] F. Zhuang, R.T. Fuchs, Z. Sun, Y. Zheng, G.B. Robb, Nucleic Acids Res. 40 (2012) e54.
[72] C.H. Jan, C.C. Williams, J.S. Weissman, Science 346 (2014) 1257521.
[73] C.C. Williams, C.H. Jan, J.S. Weissman, Science 346 (2014) 748–751.
[74] E. Oh, A.H. Becker, A. Sandikci, D. Huber, R. Chaba, F. Gloge, R.J. Nichols, A. Typas, C.A. Gross, G. Kramer, J.S. Weissman, B. Bukau, Cell 147 (2011) 1295–1308.
[75] Y. Han, A. David, B. Liu, J.G. Magadan, J.R. Bennink, J.W. Yewdell, S.B. Qian, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 12467–12472.

Figure S1

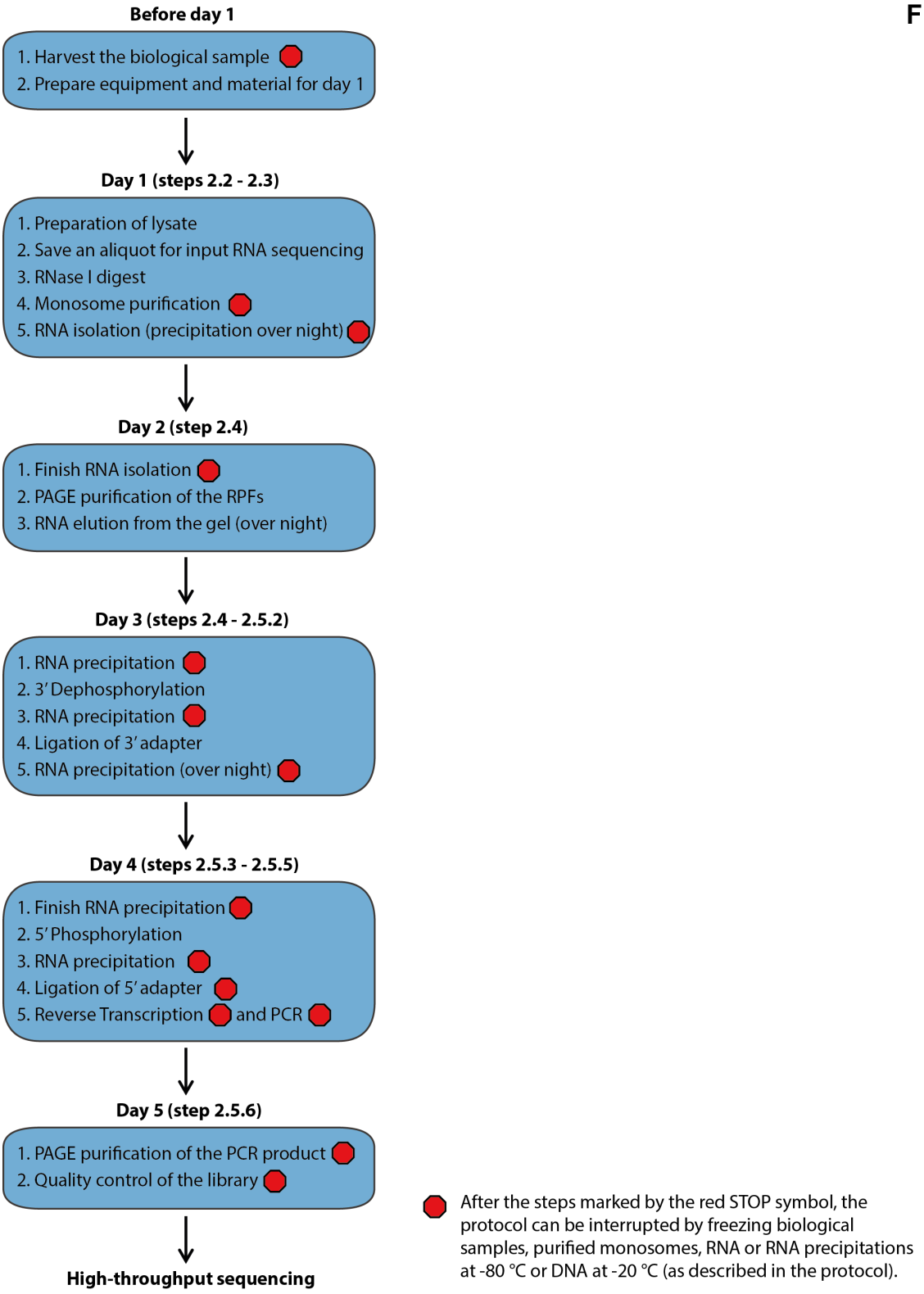


Figure S1: Flow diagram of the protocol described in this paper.
The indicated timing leads to a 5 day procedure. However, the protocol can be interrupted at various steps, as indicated by the red "STOP" symbols and described in the protocol.

Figure S2

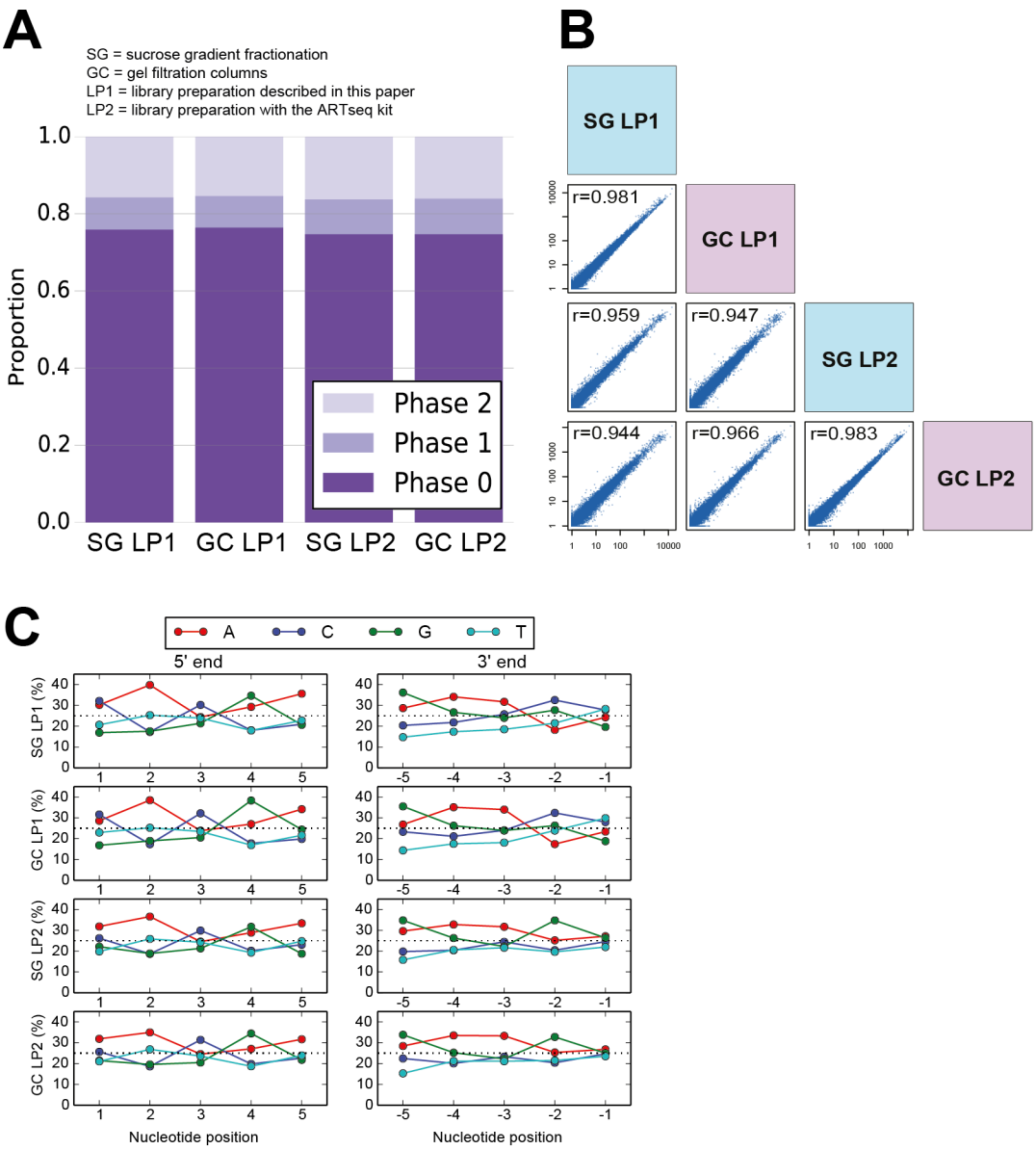


Figure S2: Analysis of the 29 nucleotides long reads in RPF libraries prepared with two different methods for isolation of monosomes and two different library preparations.

The same C. elegans lysate was subjected to two different methods of monosome purification, sucrose gradient fractionation (SG) or size exclusion chromatography using gel filtration columns (GC). RNA from both monosome samples was then prepared for sequencing using the library preparation protocol described in this paper (LP1) or the library preparation protocol of the ARTseq™ Ribosome Profiling Kit (LP2). A) Phasing of the 29 nucleotides long reads in each library. B) Pairwise correlation plot for the number of 29 nucleotides long reads mapping to the CDS of every gene on a log10 scale. C) The proportions (in percent) of each nucleotide at the first five positions from the 5' end and 3' end of all 29 nucleotides long reads mapping to coding sequence. Position "1" indicates the first position from the 5' end, whereas position "-1" indicates the first position from the 3' end.

Figure S3

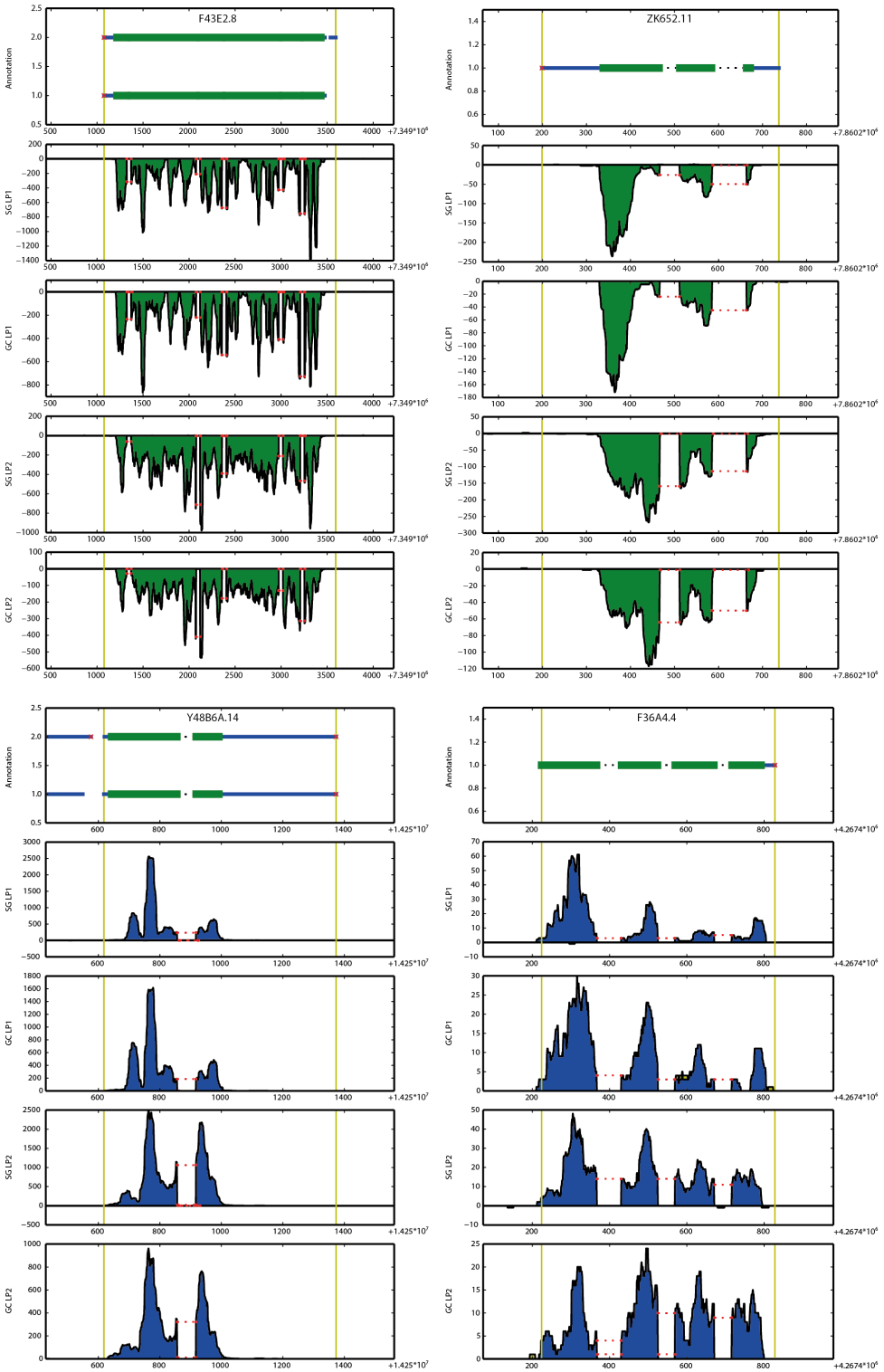


Figure S3: Read coverage of four example genes in the RPF libraries described in section 6.
The colors in the coverage figures have the following meanings: blue: unique mapped reads (+ strand); green: unique mapped reads (- strand); yellow: multiple mapped reads (both strands); red dashed line: unique mapped junction reads (both strands). In the annotation panel, the red cross indicates the 5' end of the transcripts, while thick green lines/medium blue lines/thin black dashed lines represent CDS/UTRs/introns, respectively. The x-axis shows the coordinates on the chromosome on which the respective example gene is located.

Figure S4

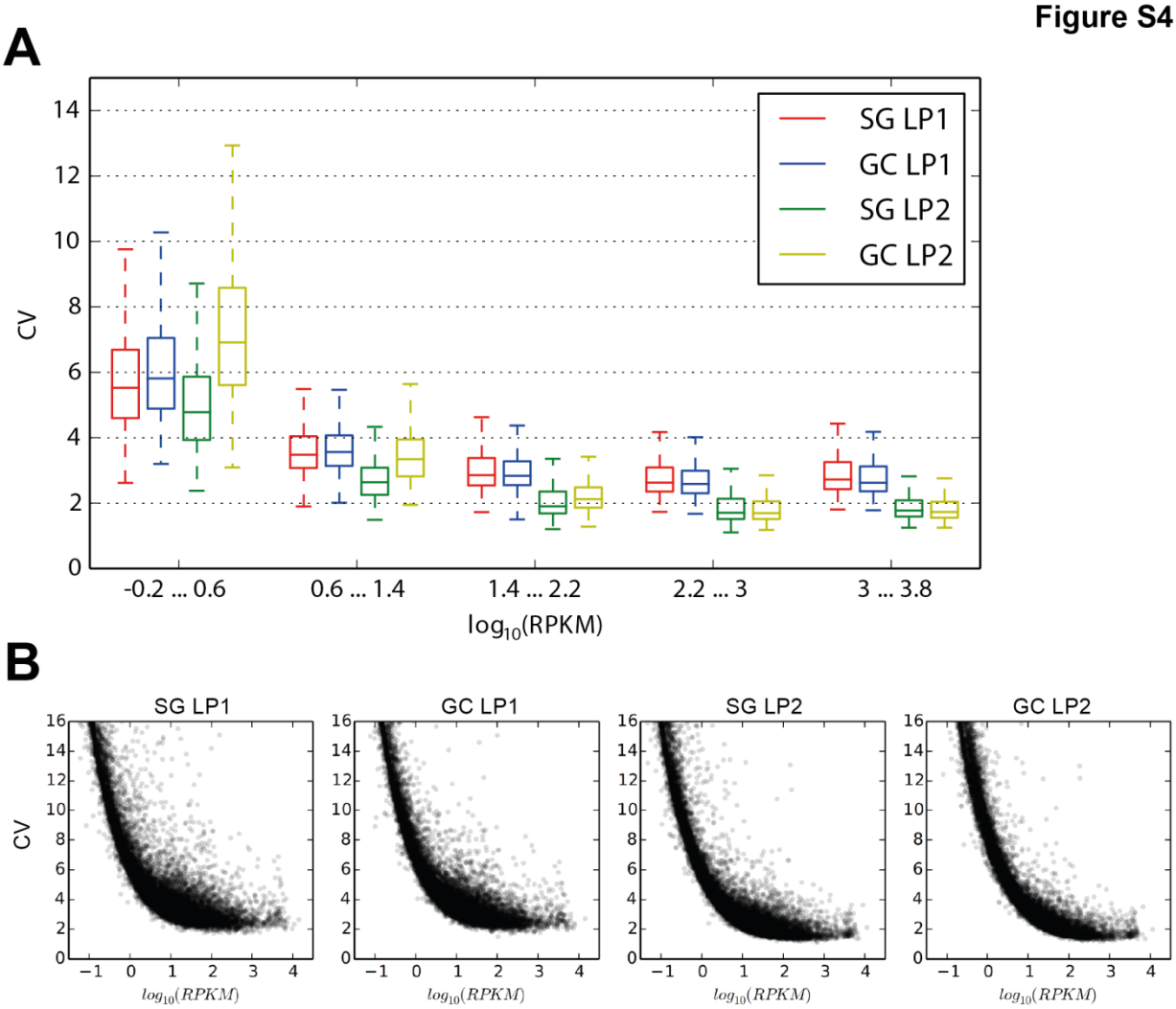


Figure S4: Coverage homogeneity of the RPF libraries described in section 6.
The read coverage (RPKM) at every nucleotide position of the CDS of each gene was determined using the transcript variant with the longest CDS and considering only the 5' terminal nucleotide of the RPFs. For the CDS of each gene, the coefficient of variation (CV) was calculated by dividing the standard deviation of read coverage per nucleotide position by the mean read coverage per nucleotide position. A) Box plots representing the distribution of CVs in groups of genes with similar expression in the different RPF libraries. For each sequenced library, genes are grouped into five categories of increasing expression according to their RPKM values (RPKM of RPFs mapping to the CDS, see Table S1 for the number of genes in each category). Outliers are not depicted. P-values for differences in CV distributions between the groups of each RPKM range were calculated using the Mann-Whitney U test and are shown in Table S3. B) For each protocol, the CV is plotted over the RPKM for each CDS.

Supplementary tables for publication:

Transcriptome-wide measurement of ribosomal occupancy by ribosome profiling

Table S1: The number of genes in the different RPKM range categories in Figure S4A for each RPF library described in section 6.

log_{10}(RPKM) range	Number of genes in SG LP1	Number of genes in GC LP1	Number of genes in SG LP2	Number of genes in GC LP2
< -0.2	8456	8473	8240	8154
-0.2 … 0.6	3970	3884	3898	3917
0.6 … 1.4	4638	4637	4766	4767
1.4 … 2.2	2646	2731	2738	2838
2.2 … 3.0	705	697	775	744
3.0 … 3.8	185	183	190	187
> 3.8	11	6	4	4

Table S2: Read numbers of the RPF libraries described in section 6.

RPF library	Number of total reads	Number of mapped reads	Number of uniquely mapped reads	Number of uniquely mapped CDS reads	Percentage of uniquely mapped CDS reads in all mapped reads
SG LP1	46005656	42681149	35017632	32494131	76.1
GC LP1	51456116	47173619	34883319	25502243	54.1
SG LP2	66437016	62697131	46558023	38132794	60.8
GC LP2	47423443	42562991	28660843	18018744	42.3

Table S3: P-values for the differences in CV distributions between the groups of each RPKM range, calculated using the Mann-Whitney U test.

log_{10}(RPKM) range	SG LP1 vs GC LP1	SG LP2 vs GC LP2	SG LP1 vs SG LP2	GC LP1 vs GC LP2
-0.2 … 0.6	$2.1896*10^{-19}$	0.0	$5.92681*10^{-125}$	$8.64387*10^{-151}$
0.6 … 1.4	$1.86473*10^{-5}$	0.0	0.0	$6.14732*10^{-59}$
1.4 … 2.2	0.189675	$2.48305*10^{-49}$	0.0	0.0
2.2 … 3.0	0.0769468	0.474607	$3.84342*10^{-144}$	$1.84627*10^{-152}$
3.0 … 3.8	0.165793	0.285673	$2.21412*10^{-39}$	$4.12147*10^{-39}$

2.2 The two modes of post-transcriptional regulation by LIN41

The establishment of the ribosome profiling technique enabled us to study the roles of the miRNA *let-7* and the RBP LIN41 during *C. elegans* larval development. We were intrigued by the finding that a failure to downregulate LIN41 protein levels in the last larval stage, in mutants disrupting the interaction of *let-7* with the *lin-41* 3'UTR, was lethal (Ecsedi et al., 2015). In developmental time course experiments, after perturbing the activities of *let-7* and/or LIN41, we studied the gene expression changes on both the level of mRNA abundance and translation, resulting in verification of direct *let-7* targets and identification of direct LIN41 targets. Thereby, we made the surprising discovery that LIN41 regulates each mRNA target with one of two distinct mechanisms, inhibition of translation initiation without affecting mRNA levels or degradation of the transcript. Moreover, we found that the choice of the applied mechanism is dependent on the location of the LIN41 binding site, with binding to the 5'UTR causing translational repression and binding to the 3'UTR causing mRNA degradation. The results of these experiments are presented in the publication of this section.

Publication

LIN41 Post-transcriptionally Silences mRNAs

by Two Distinct and Position-Dependent Mechanisms

Molecular Cell

Volume 65, February 2 (2017), Pages 476–489

doi link: http://dx.doi.org/10.1016/j.molcel.2016.12.010

Keywords:

• *lin-41*, LIN41, TRIM71, *let-7*, LIN-29, MAB-10, MAB-3, DMD-3, EGR, NAB

• RNA binding protein

• posttranscriptional regulation, translational repression, mRNA degradation

• positional effect, 3'UTR, 5'UTR

• self-renewal, stem cell

• ribosome profiling

• *Caenorhabditis elegans*, seam cell, heterochronic

Highlights:

• Ribosome profiling and RNA immunoprecipitation identify direct targets of LIN41.

• Three mRNAs undergo LIN41-dependent decay, one undergoes pure translational repression.

• The location of LIN41 binding on the mRNA determines the choice of mechanism.

• LIN-29A/EGR and MAB-10/NAB may be evolutionarily conserved LIN41 targets.

Comments:

• Tables S1 and S2 are not included in this thesis.

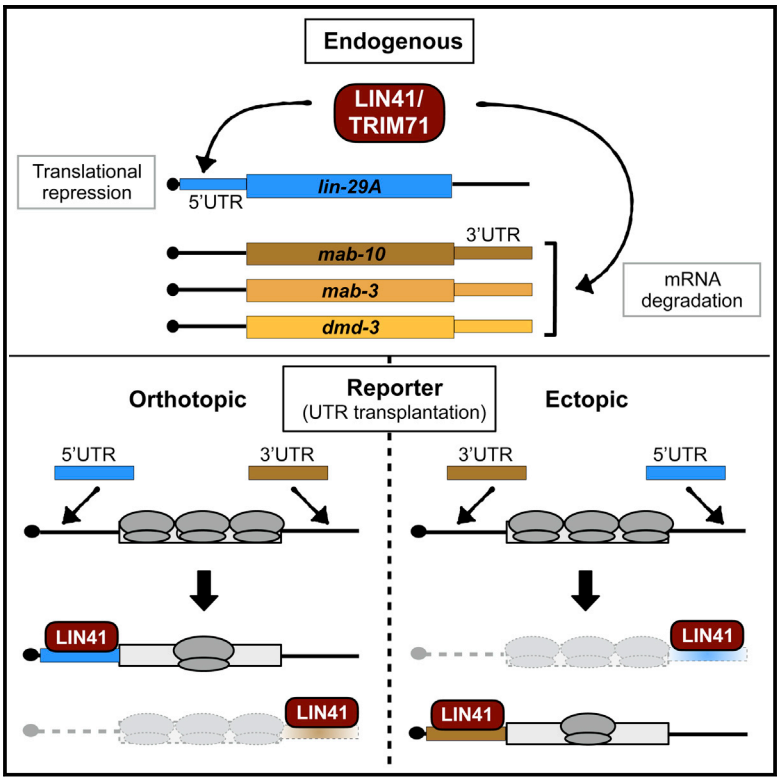
• Tables S3 and S5 are attached to the end of the publication.

Article

Molecular Cell

LIN41 Post-transcriptionally Silences mRNAs by Two Distinct and Position-Dependent Mechanisms

Graphical Abstract



Highlights

- Ribosome profiling and RNA immunoprecipitation identify direct targets of LIN41
- Three mRNAs undergo LIN41-dependent decay, one undergoes pure translational repression
- The location of LIN41 binding on the mRNA determines the choice of mechanism
- LIN-29A/EGR and MAB-10/NAB may be evolutionarily conserved LIN41 targets

Authors

Florian Aeschimann, Pooja Kumari, Hrishikesh Bartake, Dimos Gaidatzis, Lan Xu, Rafal Ciosk, Helge Großhans

Correspondence

helge.grosshans@fmi.ch

In Brief

Aeschimann et al. report that the stem cell fate regulator LIN41 silences mRNAs through two distinct and separable mechanisms, translational repression or degradation. Choice of mechanism is specified by the location of the LIN41-binding site on the target mRNA.

Aeschimann et al., 2017, Molecular Cell *65*, 476–489
February 2, 2017
http://dx.doi.org/10.1016/j.molcel.2016.12.010

CellPress

Molecular Cell
Article

LIN41 Post-transcriptionally Silences mRNAs by Two Distinct and Position-Dependent Mechanisms

Florian Aeschimann,[1,2] Pooja Kumari,[1] Hrishikesh Bartake,[1,2] Dimos Gaidatzis,[1,3] Lan Xu,[1] Rafal Ciosk,[1] and Helge Großhans[1,4,*]

[1]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland
[2]University of Basel, Faculty of Science, 4056 Basel, Switzerland
[3]Swiss Institute of Bioinformatics, 4058 Basel, Switzerland
[4]Lead Contact
*Correspondence: helge.grosshans@fmi.ch
http://dx.doi.org/10.1016/j.molcel.2016.12.010

SUMMARY

The RNA-binding protein (RBP) LIN41, also known as LIN-41 or TRIM71, is a key regulator of animal development, but its physiological targets and molecular mechanism of action are largely elusive. Here we find that this RBP has two distinct mRNA-silencing activities. Using genome-wide ribosome profiling, RNA immunoprecipitation, and in vitro-binding experiments, we identify four mRNAs, each encoding a transcription factor or cofactor, as direct physiological targets of *C. elegans* LIN41. LIN41 silences three of these targets through their 3′ UTRs, but it achieves isoform-specific silencing of one target, *lin-29A*, through its unique 5′ UTR. Whereas the 3′ UTR targets *mab-10*, *mab-3*, and *dmd-3* undergo transcript degradation, *lin-29A* experiences translational repression. Through binding site transplantation experiments, we demonstrate that it is the location of the LIN41-binding site that specifies the silencing mechanism. Such position-dependent dual activity may, when studied more systematically, emerge as a feature shared by other RBPs.

INTRODUCTION

Proper formation and homeostasis of tissues and organs requires switching of stem and progenitor cells from self-renewal to an appropriate differentiation program in the right place and at the correct time. Post-transcriptional mechanisms, although less well studied than contributions of transcriptional control, have been argued to play a dominant role in regulating stem cell fates (Wright and Ciosk, 2013; Ye and Blelloch, 2014). LIN41 and its regulator, the microRNA (miRNA) *let-7*, appear to have major and conserved functions in these processes. They control proliferation versus differentiation decisions not only in *C. elegans* seam cells (Reinhart et al., 2000; Slack et al., 2000), epidermal blast cells considered an in vivo stem cell model (Brabin and Woollard, 2012; Joshi et al., 2010), but also in mammalian embryonic stem cells and during human fibroblast in vitro reprogramming (Chang et al., 2012; Chiu et al., 2014; Rehfeld et al., 2015; Rybak et al., 2009; Worringer et al., 2014). It has therefore been speculated that the regulation of LIN41 by *let-7* constitutes an ancient control mechanism for self-renewal, differentiation, and cell fate plasticity in diverse tissues (Ecsedi and Großhans, 2013). Moreover, LIN41 is the one key target of *let-7* in *C. elegans* whose regulation ensures proper vulval development and, thus, viability (Ecsedi et al., 2015).

The molecular mechanisms by which LIN41 (also called TRIM71 in mammals) exerts its functions are not well understood. As a member of the TRIM-NHL protein family, post-transcriptional or post-translational mechanisms are likely (Tocchini and Ciosk, 2015). This is because the eponymous *tri*partite *m*otif of RING, B-Box, and coiled-coil domains is characteristic of proteins with E3 ubiquitin ligase activity (Ikeda and Inoue, 2012), whereas the C-terminal NHL (NCL-1, HT2A2, and LIN-41) repeat domain may mediate sequence-specific RNA binding (Loedige et al., 2015). A protein ubiquitylation activity has been established for mouse LIN41 in some contexts (Chen et al., 2012; Rybak et al., 2009), but, as *C. elegans* and *D. melanogaster* LIN41 may lack this activity (Löer et al., 2008; Tocchini et al., 2014), this seems not to account for a mechanistically conserved self-renewal activity across animal phylogeny. By contrast, LIN41 may have a conserved function in mRNA silencing. A role for LIN41 in translational repression of mRNAs was first proposed more than 15 years ago in *C. elegans* (Slack et al., 2000) and since then has been suggested repeatedly in diverse systems (Loedige et al., 2013; Spike et al., 2014b; Worringer et al., 2014). However, this notion has not been tested explicitly, but it was deduced from the observation that certain LIN41 target reporters change more extensively at the level of reporter protein activity than mRNA level (Loedige et al., 2013). Indeed, extensive evidence supports a function of LIN41 in destabilizing target mRNAs (Chang et al., 2012; Loedige et al., 2013; Mitschka et al., 2015), even in cases where this RNA-binding protein (RBP) was concluded to act by translational repression (Worringer et al., 2014).

Progress toward understanding the mode of action of LIN41 has suffered from limited knowledge of physiological LIN41 targets. In *C. elegans*, genetic interactions link CDC-25.3, a meiotic regulator (Spike et al., 2014a), and LIN-29, a transcription factor that regulates seam cell self-renewal and differentiation (Ambros and Horvitz, 1984; Rougvie and Ambros, 1995; Slack et al.,

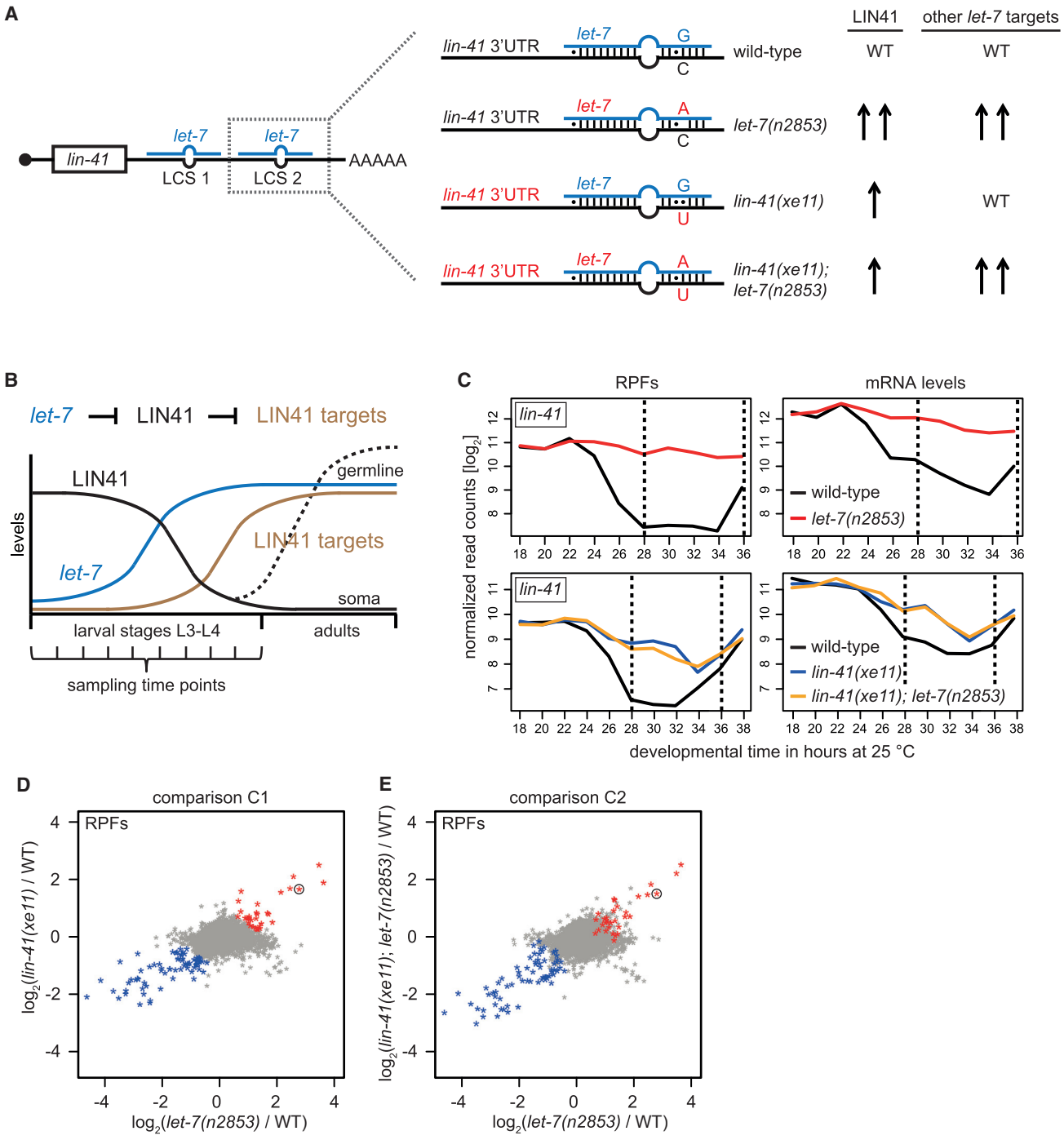


Figure 1. Gene Expression Changes Caused by Dysregulation of LIN41

(A) Schematic of *let-7* miRNA binding to the two functional *let-7* complementary sites (LCSs) in the *lin-41* 3′ UTR (Vella et al., 2004). Blow-ups illustrate the effects of mutations for only the second LCS. Lines indicate Watson-Crick base pairs, dots represent wobble base pairs, and mutated genes and nucleotides are in red. Columns on the right illustrate how the levels of LIN41 and of other *let-7* targets are affected (WT, wild-type levels; one arrow, partial de-silencing; two arrows, full de-silencing).

(B) Schematic of developmental expression patterns of LIN41, its regulator *let-7*, and its hypothetical targets. Following silencing of LIN41 in the soma by *let-7*, LIN41 accumulates massively in adult germlines. Synchronized worm populations were harvested bi-hourly as indicated.

(C) *lin-41* expression over time at the level of ribosome-protected fragments (RPFs, left) and mRNA (right). Upper and lower panels represent two independent experiments. Fewer reads obtained from the second time course experiment account for generally lower normalized $\log_2$ read counts relative to the first time course experiment. The two dashed lines indicate the window of time points pooled for differential gene expression analysis in (D) and (E).

(legend continued on next page)

2000), to LIN41 functions in the adult germline and larval epidermis, respectively. However, both await experimental validation as direct LIN41 targets.

Here we identify direct targets of *C. elegans* LIN41. These include *lin-29A/EGR* and *mab-10/NAB*, implicated in LIN41-dependent mammalian cell fate reprogramming (Worringer et al., 2014), implying evolutionary conservation of a LIN41-dependent fundamental stem cell fate regulatory process. We find that LIN41 can silence mRNAs through two distinct mechanisms, repression of translation or destabilization. Unexpectedly, the choice of mechanism depends on the target and is instructed by where on the mRNA LIN41 binds: binding to the 5′ UTR elicits translational repression, and binding to the 3′ UTR elicits transcript degradation. We are currently aware of two additional examples of animal RBPs with position-dependent dual activities (Beckmann et al., 2005; Kühn, 2015), each with unique combinations of activities. Therefore, more such RBPs may remain to be discovered.

RESULTS

Dysregulation of LIN41 Quantitatively Explains Gene Expression Changes in *let-7* Mutant Animals

To elucidate the regulatory functions of LIN41, we compared changes in gene expression between wild-type animals and previously described *C. elegans* mutants, in which LIN41 is uncoupled from repression by *let-7* to different extents as follows (Figure 1A): (1) *let-7*(*n2853*) mutant animals carry a point mutation in the *let-7* seed sequence (Reinhart et al., 2000) that prevents *let-7* activity at 25°C, resulting in complete de-silencing of *lin-41* and the other *let-7* targets. Thus, a comparison to wild-type animals identifies any gene dysregulated in the absence of *let-7*, be it by direct *let-7* targeting or as a secondary effect. (2) *lin-41*(*xe11*) mutant animals contain two point mutations in the *lin-41* 3′ UTR, one in each of the two functional *let-7* complementary sites (LCSs) (Ecsedi et al., 2015). The resulting replacement of a G:C Watson-Crick base pair by a G:U wobble pair in the two seed:seed-match hybrids formed with wild-type *let-7* causes partial but specific de-silencing of *lin-41*. (3) *lin-41*(*xe11*); *let-7*(*n2853*) double-mutant animals carry the point mutations of both (1) and (2). Hence, the mutation in the let-7 seed sequence disrupts silencing of all *let-7* targets but *lin-41*, whose two LCSs each contain a compensatory mutation in the seed match that restores base pairing. However, *let-7* levels in the *let-7*(*n2853*) genetic background are reduced (Chatterjee and Großhans, 2009; Reinhart et al., 2000), preventing a full, wild-type-like repression of *lin-41* in this situation. Thus, *lin-41*(*xe11*); *let-7*(*n2853*) double-mutant animals exhibit a similar partial de-silencing of *lin-41* as the *lin-41*(*xe11*) single-mutant animals (see below and Ecsedi et al., 2015), but full de-silencing of all other *let-7* targets.

To identify transcripts that LIN41 might regulate through translational repression or degradation, we performed ribosome profiling and RNA sequencing (RNA-seq) on synchronized worm populations, sampled every 2 hr during development from late larval stage 2 (L2)/early L3 to late L4/young adult stages (Figure 1B; Tables S1 and S2 provide normalized $\log_2$ read counts). These time course experiments offered two advantages over single time point measurements. First, because *let-7* levels increase greatly between L3 and L4 stages (Reinhart et al., 2000), *lin-41* is presumably increasingly repressed in this time window (Figure 1B). Since LIN41 was suggested to be an RBP with repressive function, we predicted LIN41 downregulation to cause an accumulation of its targets over time. Second, as gene expression in *C. elegans* is highly dynamic, with thousands of genes exhibiting rhythmic expression with high amplitude (Hendriks et al., 2014), single time point experiments may be prone to expression artifacts through differences in developmental rates (and thus misalignment of time points) and/or population synchrony between wild-type and mutant animals (Figure S1).

We compared wild-type to *let-7*(*n2853*) animals in a first experiment, and we compared a biological wild-type replicate to both *lin-41*(*xe11*) single- and *lin-41*(*xe11*); *let-7*(*n2853*) double-mutant animals in a second experiment. In wild-type worms, both *lin-41* mRNA and ribosome-protected fragment (RPF) levels started to decrease from late L3/early L4 stage on, corresponding to the time of increase in *let-7* expression (Figure 1C). They reached a plateau by early/mid-L4 before rising again in late L4 stage, when *lin-41* starts being expressed in the germline (Spike et al., 2014a; Tocchini et al., 2014) (data not shown). As observed previously (Bagga et al., 2005; Ding and Großhans, 2009), major decreases were apparent at the level of the transcript, but they appeared somewhat enhanced at the translational level. In *let-7*(*n2853*) mutant animals, both types of repression were completely eliminated, and *lin-41* mRNA and RPF levels remained at L3 level throughout L4. In *lin-41*(*xe11*) single- and *lin-41*(*xe11*); *let-7*(*n2853*) double-mutant animals, repression of *lin-41* occurred but was blunted relative to wild-type.

Next, we examined gene expression changes between the mutants and their corresponding wild-type controls, averaged from 28 to 36 hr of development, the time window of the *lin-41* repression plateau (Figure 1C). We focused on fold changes at the level of RPFs, as these would integrate RNA level and translational changes, and we performed two comparisons, referred to as "C1" and "C2." In C1, we examined the effect of fully dysregulating all *let-7* targets (in *let-7*(*n2853*) animals) versus partially dysregulating only LIN41 (in *lin-41*(*xe11*) animals) (Figure 1D). This revealed substantial similarity in the genes dysregulated in each mutant relative to wild-type. In agreement with higher levels of LIN41 in the *let-7*(*n2853*) than in the *lin-41*(*xe11*) mutant background, the extent of dysregulation of individual genes was consistently larger in the *let-7* mutant animals. Taking this into account, we identified sets of genes consistently up- or downregulated in the two mutants (Figure 1D, red and blue

(D and E) Scatterplots depicting mutant to wild-type $\log_2$ fold changes in normalized RPF read counts for each gene. Gene expression changes are compared between (D) *let-7*(*n2853*) and *lin-41*(*xe11*) and (E) *let-7*(*n2853*) and *lin-41*(*xe11*); *let-7*(*n2853*) mutant animals. In (D) and (E), genes upregulated in both *let-7*(*n2853*) and *lin-41*(*xe11*) mutants are colored red, those downregulated are colored blue (METHODS), and *lin-41* is circled. Each comparison (x axis versus y axis) is between two independent experiments, with independent wild-type replicates.
See also Figure S1 and Tables S1 and S2.

asterisks, respectively), i.e., genes changed upon LIN41 dysregulation.

The similarity of gene expression changes caused by the two distinct mutations in C1 suggested that LIN41 upregulation accounted for many of the gene expression changes in *let-7*(*n2853*) animals. To test this notion, we examined, in C2, the effect of fully dysregulating all *let-7* targets in a context of complete (in *let-7*(*n2853*) animals) or partial (in *lin-41*(*xe11*); *let-7*(*n2853*) double-mutant animals) LIN41 de-silencing (Figure 1E). As expected from the shared *let-7* mutation, gene expression changes overlapped extensively. Strikingly, however, when we highlighted the genes that were consistently up- or downregulated in C1, these were largely identical to those consistently dysregulated in C2. Only a few additional upregulated genes emerged that had not been upregulated in *lin-41*(*xe11*) in C1 (gray asterisks in upper right quadrant of Figure 1E). These included direct *let-7* targets such as *daf-12* (Großhans et al., 2005) and *hbl-1* (Abrahante et al., 2003) (Figures S2A and S2B). Hence, although the *let-7*(*n2853*) mutation de-silences *let-7* targets broadly, secondary changes largely depend on dysregulation of *lin-41*. Indeed, the magnitude of gene expression changes appeared proportionate to the extent of *lin-41* de-silencing, i.e., it was greater in *let-7*(*n2853*) than in *lin-41*(*xe11*) or *lin-41*(*xe11*); *let-7*(*n2853*) mutant animals. We conclude that *let-7* effects on gene expression are extensively and quantitatively explained by dysregulation of LIN41 as its primary target.

Identification of Direct LIN41 Target Genes

To identify direct targets of LIN41, we visually examined temporal changes in gene expression for the selected genes changed upon LIN41 dysregulation, reasoning that effects on the levels of primary targets should precede those on the levels of secondary targets. The first gene that consistently changed in the various mutants relative to wild-type animals was *lin-41*, whose expression was increased in all mutants relative to wild-type by 24 hr (Figures 1C and 2A; Table S3). For the six genes affected next, termed "mid" in Figure 2A, RPF levels were all decreased in the mutants relative to wild-type. Specifically, their levels all increased in L4-stage wild-type animals as LIN41 disappeared, but they remained low in L4-stage *let-7*(*n2853*) or *lin-41*(*xe11*) mutant animals, which retain high LIN41 levels (Figures 2B and S2C). This is the pattern we would predict for direct targets of LIN41, and the group included *lin-29*, a previously proposed target of LIN41 (Slack et al., 2000). Subsequently affected genes were either up- or downregulated, consistent with secondary effects dominating in this class. Closer inspection of the expression patterns of the mid genes revealed that, for five of the six genes, both transcript and RPF levels changed (Figures 2B and S2C), suggesting potential regulation at the level of transcript stability. *lin-29* was an exception, where little or no difference occurred on the mRNA level between wild-type and mutant animals, but major differences occurred on the RPF level (Figure 2B). This finding suggested a predominant or exclusive regulation of *lin-29* at the translational level.

To test if the six mid genes are directly regulated by LIN41, we sought to determine whether LIN41 physically interacted with their mRNAs in L3/L4-stage animals. By immunoprecipitation (IP) of a functional FLAG::GFP::LIN41 fusion protein, we detected binding of LIN41 to *lin-29* and *mab-10* mRNAs as well as, to a lesser extent, *mab-3* and *dmd-3* mRNAs (Figure 2C). In contrast, the other two candidate targets, *ceh-60* and *Y54G2A.3*, behaved similarly to the negative control mRNAs, *act-1* and *unc-54*. Binding of *lin-29*, *mab-10*, *mab-3*, and *dmd-3* mRNAs was specific in that it was not observed with IP of another RBP, FLAG::GFP::SART-3 (Rüegger et al., 2015) (Figure 2C). We conclude that these four genes may be direct targets of LIN41.

LIN41-Dependent Regulation through Target Gene 3′ UTRs

RBPs frequently confer regulation by binding to the 3′ UTRs of target mRNAs, and LIN41 is capable of repressing target reporter genes in this manner in mammalian cells in vitro (Chang et al., 2012; Loedige et al., 2013). To determine whether LIN41 exerts a similar function in vivo in *C. elegans*, we constructed reporter transgenes consisting of the ubiquitously and constitutively expressing *dpy-30* promoter, a sequence encoding GFP fused to a destabilizing PEST sequence and nuclear H2B to achieve greater expression dynamics and nuclear concentration of the signal, respectively, and a 3′ UTR of interest (Figure 3). All transgenes were integrated in single copy into the same genomic location, and expression was examined by confocal microscopy in the epidermis of L3-stage animals, i.e., prior to LIN41 repression by *let-7*. For each of the following GFP reporter experiments, the GFP signals for at least ten worms were observed to verify that they were comparable among different worms in each transgenic line and for each condition.

As a control, we utilized the *unc-54* 3′ UTR, not known to confer any post-transcriptional regulation. As expected, GFP signal was readily detectable in the epidermis for this reporter, irrespective of the presence of LIN41 (Figure 3). By contrast, a reporter containing the *mab-10* 3′ UTR was silenced extensively in the same tissue. Silencing was dependent on LIN41, as RNAi-mediated depletion of LIN41 relieved it. Use of the *mab-3* and *dmd-3* 3′ UTR yielded similar results. Surprisingly, however, we observed no significant repressive activity of LIN41 on the *lin-29* 3′ UTR. Taken together, these and the above findings establish *mab-3*, *mab-10*, and *dmd-3* as bona fide LIN41 targets, whose regulation involves transcript degradation conferred by their 3′ UTRs.

LIN41 Regulates Only the A Isoform of *lin-29*

To understand *lin-29* translational regulation, we inspected its RPF profiles in more detail. Inhibition of translation initiation would lead to a uniform decrease in RPF coverage along the *lin-29* open reading frame (ORF), while inhibition of translation elongation or premature ribosome drop-off could lead to a decrease in RPF coverage toward the 3′ end of the ORF. To look at changes in the RPF distribution in the *let-7* mutant compared to wild-type, we calculated the fold repression on a per-exon basis, summing up the five time points of the *lin-41* repression plateau (28–36 hr, see Figure 1C). As a control, we performed the same analysis for mRNA reads. Unexpectedly, we observed that the sustained *lin-41* expression in *let-7* mutant animals caused an apparently greater translational repression toward the *lin-29* 5′ end than the 3′ end, with an ~10-fold

A

change in time course	genes up	genes down
early	1*	0
mid	0	6**
late	24	64

*lin-41 **ceh-60
mab-10
lin-29
dmd-3
mab-3
Y54G2A.3

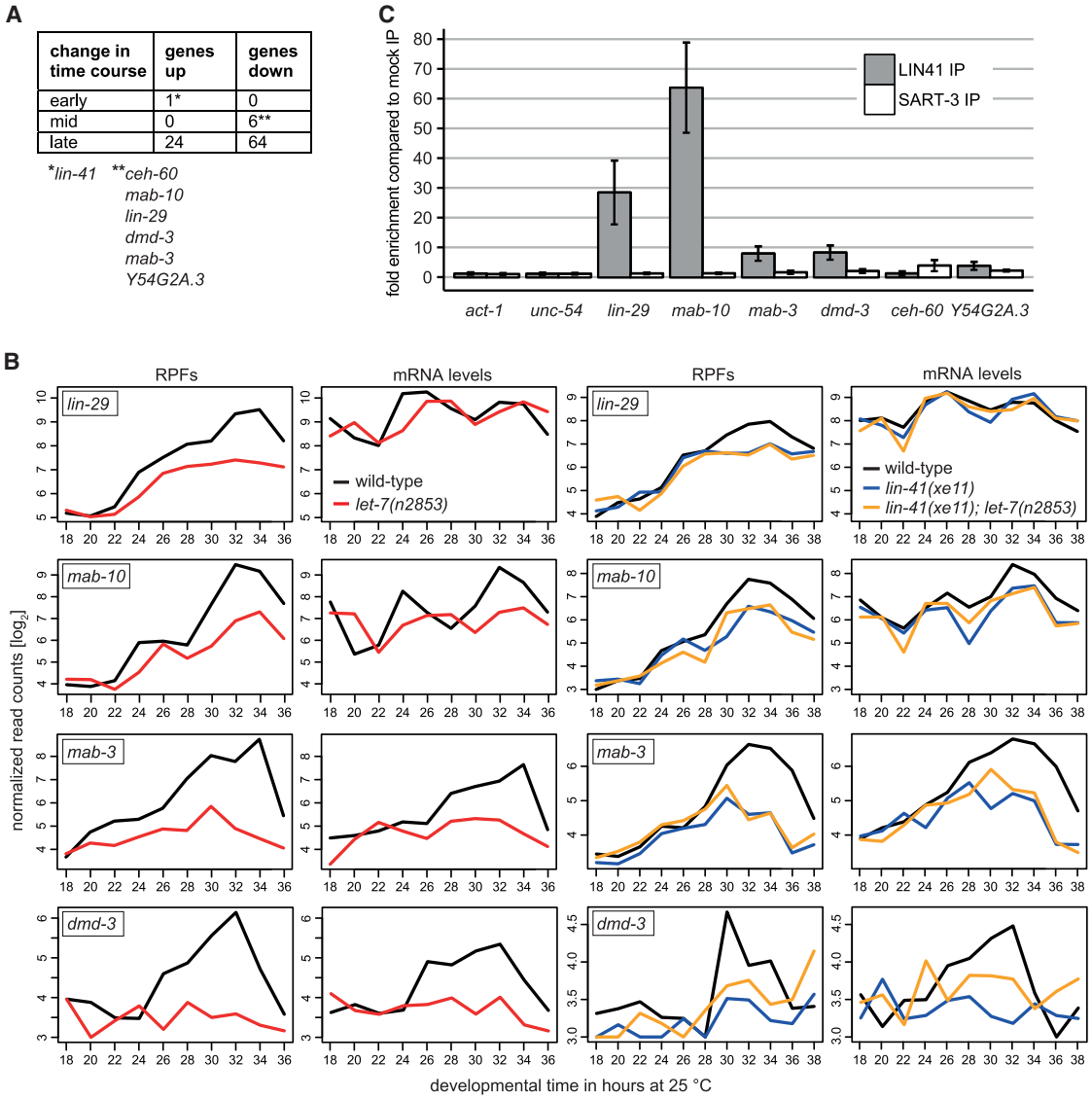


Figure 2. Identification of LIN41 Target mRNAs

(A) Genes differentially expressed in all examined mutant (*let-7*(*n2853*), *lin-41*(*xe11*), and *lin-41*(*xe11*); *let-7*(*n2853*)) relative to wild-type animals are listed according to the earliest time point of consistent dysregulation at the RPF level. "Up" and "down" refer to expression changes in mutants relative to wild-type, not to expression changes over time. Table S3 provides a complete list and details.

(B) Expression of candidate LIN41 target genes over development is shown.

(C) RT-qPCR analysis on RNA co-immunoprecipitated with FLAG::GFP::LIN41 or an unrelated RBP, FLAG::GFP::SART-3. mRNA levels of six candidate LIN41 targets and the negative controls *act-1* and *unc-54* were determined. Immunoprecipitation (IP) occurred through an anti-FLAG antibody, and fold enrichments were calculated relative to anti-FLAG IP in non-transgenic (wild-type) animals. Worms were harvested as semi-synchronous L3/L4-stage animals. n = 4 biological replicates, data as mean ± SEM.

See also Figure S2 and Table S3.

repression for each of the exons one through four but only an ~2-fold repression for each of the other exons (Figure 4A). In contrast to RPF levels, mRNA levels were unaffected for any exon.

We reasoned that, rather than pointing to a specific mechanism of translational repression, the difference between exons might reflect differential regulation of *lin-29* isoforms, as the two reported *lin-29* isoforms encompass exons 1–11 (long isoform *lin-29A*) and 5–11 (short isoform *lin-29B*), respectively (Figure 4A) (Rougvie and Ambros, 1995). A preferential regulation of *lin-29A* by LIN41 would explain why exons 1–4, which are exclusive to *lin-29A*, are more strongly regulated than exons 5–11, which are shared by the A and B isoforms. In other words, the change in RPF reads on exons 5–11 caused by regulation of *lin-29A* might be partially masked by reads from the unregulated *lin-29B* isoform. The difference between exons 1–4 and exons

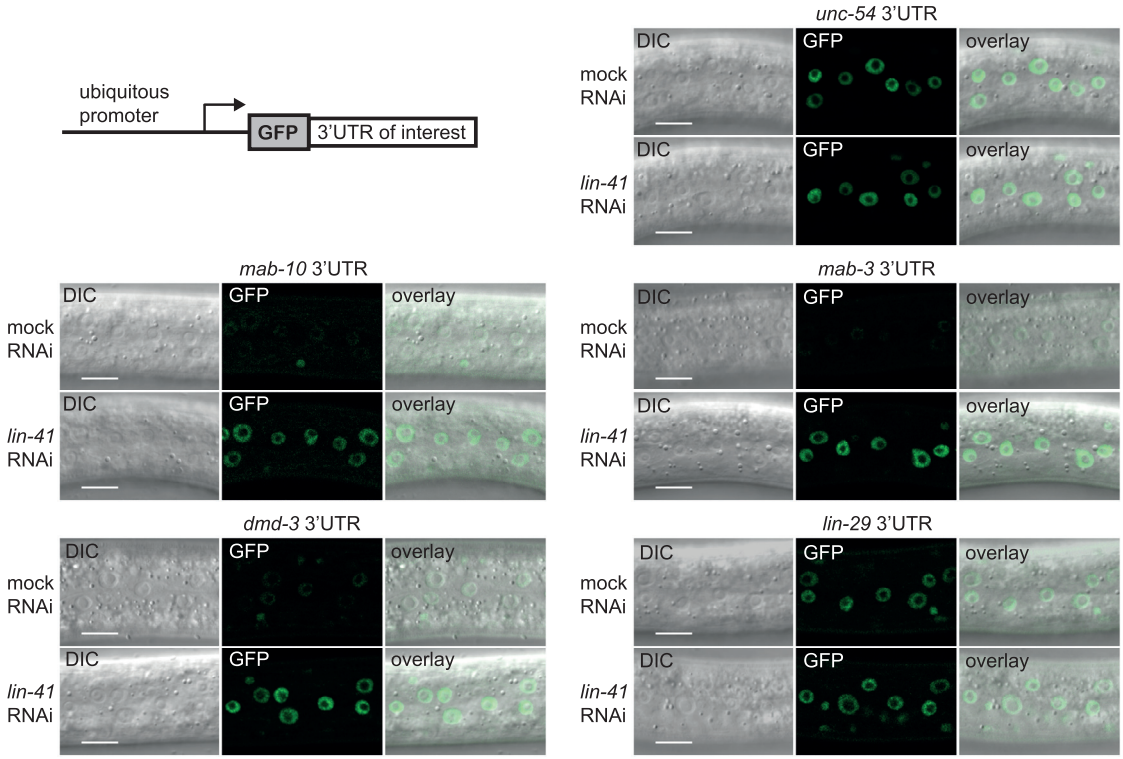


Figure 3. The 3′ UTRs of *mab-10*, *mab-3*, and *dmd-3* mRNAs Confer LIN41-Dependent Gene Silencing

The constitutive *dpy-30* promoter drives ubiquitous expression of mRNA encoding a nuclear-localized fluorophore (GFP(PEST)::H2B; labeled "GFP") and containing to the 3′ UTRs of candidate LIN41 targets or of *unc-54* as a control. Animals were grown on *lin-41* RNAi or mock RNAi bacteria, and images of epidermal nuclei of live early L3-stage worms were acquired by confocal imaging. Shown are images with differential interference contrast (DIC), GFP (identical settings for both RNAi conditions), and the overlay of the two. Scale bars here and in all other figures, 10 μm.

5–11 also held true when examining the effect on *lin-29* translation over time (Figure 4B) and when performing the same analyses with the *lin-41*(*xe11*) and *lin-41*(*xe11*); *let-7*(*n2853*) mutant animals (Figures S3A and S3B). To verify preferential regulation of *lin-29A* by LIN41, we examined the extent to which isoforms co-immunoprecipitated with LIN41. Consistent with preferential binding to *lin-29A*, LIN41 IP enriched this mRNA isoform, but not the shorter *lin-29B* isoform (Figures 4C and S3C).

To confirm the differential effect of LIN41 on LIN-29 isoforms directly on the protein level, we used genome editing to place a GFP::3xFLAG tag on the shared C terminus of endogenous LIN-29A and LIN-29B. These animals appear overtly wild-type, indicating functionality of the tagged protein, and they recapitulate temporal and spatial expression patterns previously established by immunofluorescence (Bettinger et al., 1996, 1997; data not shown). Moreover, and consistent with LIN41-mediated silencing, LIN-29::GFP protein was undetectable in the epidermis of live L3-stage animals exposed to mock RNAi, but it revealed strong nuclear accumulation upon LIN41 depletion (Figure 4D). By contrast, L3-stage nuclear accumulation of LIN-29 occurred in some non-epidermal tissues, such as the pharynx, independently of LIN41 depletion.

To test for differential regulation of the two LIN-29 protein isoforms, we performed western blotting. This revealed that, in L3-stage animals and thus in the presence of LIN41, LIN-29B, but not LIN-29A, was detectable in total animal lysates (Figures 4E and S3D). Moreover, little or no change occurred for LIN-29B when animals were depleted of LIN41. In striking contrast, LIN41 depletion caused a strong accumulation of LIN-29A protein. We conclude that LIN41 preferentially or exclusively regulates isoform A of LIN-29. As depletion of LIN41 left *lin-29A* mRNA levels unaffected (Figures 4F, S3E, and S3F), this regulation occurred on the level of translation, as expected from the ribosome profiling experiment.

Silencing of *lin-29A* Occurs through Its 5′ UTR

Although the *lin-29* 3′ UTR had no repressive activity (Figure 3), we were able to recapitulate the strong endogenous *lin-29A* regulation with a reporter construct, when we placed GFP(PEST)::H2B between the 4-kb region upstream of the *lin-29A* start codon and the *lin-29A* 3′ UTR (Figure 5A). The 4-kb region contained both the putative *lin-29A* promoter and 5′ UTR, but regulation appeared unlikely to involve the promoter since modulation of LIN41 left endogenous *lin-29A* mRNA levels unaffected. Therefore, we examined the consequences of replacing the first exon of the *lin-29A* 5′ UTR with an *act-1* 5′ UTR exon. This caused a loss of repression, in contrast to exchange of the *lin-29* 3′ UTR by the *unc-54* 3′ UTR (Figure 5A). Hence, LIN41-dependent silencing requires an intact *lin-29A* 5′ UTR.

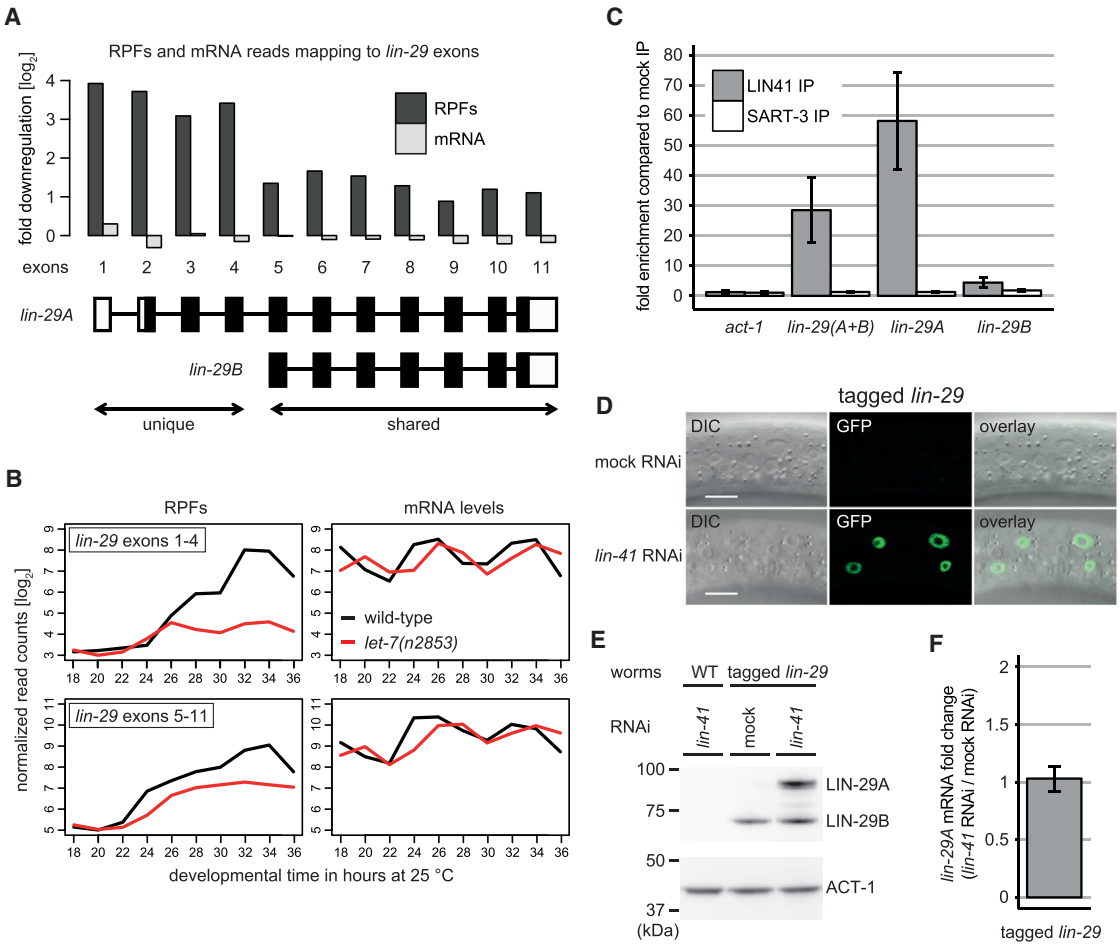


Figure 4. LIN41 Inhibits the Translation of Only One of the Two *lin-29* Isoforms

(A) Quantification of the fold downregulation in RPF and mRNA reads (log_2) in *let-7* mutant relative to wild-type worms for each exon of the *lin-29* gene. Reads were pooled from the five time points used for differential gene expression analysis in Figures 1D and 1E (28–36 hr). Below the histogram, a schematic representation (not to scale) depicts *lin-29* isoforms (open boxes, UTRs; filled boxes, coding sequence).

(B) Expression of *lin-29* over development, separated by reads pooled from exons 1–4 (unique to *lin-29A*) and exons 5–11 (shared between *lin-29A* and *lin-29B*), is shown.

(C) RT-qPCR analysis to test for enrichment of the two *lin-29* isoform mRNAs by LIN41 coIP, as described for Figure 2C. *act-1* mRNA is a negative control. The unique SL1 splice leader sites were exploited to distinguish between the two *lin-29* isoforms (see Figure S3C). n = 4 biological replicates, data as mean ± SEM.

(D) Confocal images show the endogenously tagged LIN-29 protein, accumulating in epidermal nuclei of early L3-stage *lin-29*(*xe61*[*lin-29*::*gfp*::*3xflag*]) worms upon depletion of LIN41.

(E) Western blot analysis to detect endogenous GFP::3xFLAG-tagged LIN-29A and LIN-29B proteins in early L3-stage *lin-29*(*xe61*[*lin-29*::*gfp*::*3xflag*]) animals using an anti-FLAG antibody. ACT-1 was detected as a loading control. Wild-type (WT) worms without the inserted GFP::3xFLAG tag control for antibody specificity.

(F) RT-qPCR analysis to measure the fold change of *lin-29A* mRNA levels (normalized by *act-1* mRNA levels) in early L3-stage *lin-29*(*xe61*[*lin-29*::*gfp*::*3xflag*]) animals exposed to *lin-41* RNAi relative to mock RNAi is shown. n = 3 biological replicates, data as mean ± SEM.

See also Figure S3.

To test whether the *lin-29A* 5′ UTR was sufficient for LIN41-mediated silencing, we placed it upstream of GFP and the *unc-54* 3′ UTR in a reporter expressed from the constitutively active *dpy-30* promoter. The *lin-29A* 5′ UTR comprises exon 1 and part of exon 2. Therefore, we generated two reporters that either included the 5′ UTR sequences from both exons and the intron or only exon 1. Both reporters were silenced by LIN41 (Figure 5B). By contrast, constructs that contained either the complete sequence or only exon 1 of the *mab-10* 5′ UTR were not detectably regulated (Figures 5B and S4A). We conclude that exon 1 of the 5′ UTR is both necessary and sufficient for *lin-29A* silencing through LIN41 and that repression is independent of 5′ UTR splicing.

LIN41 Can Bind Directly to Its Target mRNAs

The NHL domain of the TRIM-NHL proteins Brat and LIN41 can bind to RNA (Kwon et al., 2013; Loedige et al., 2014, 2015). To test whether LIN41 may contact its mRNA targets directly via its NHL domain, we expressed a recombinant LIN41 variant consisting of the C-terminal Filamin and NHL domains

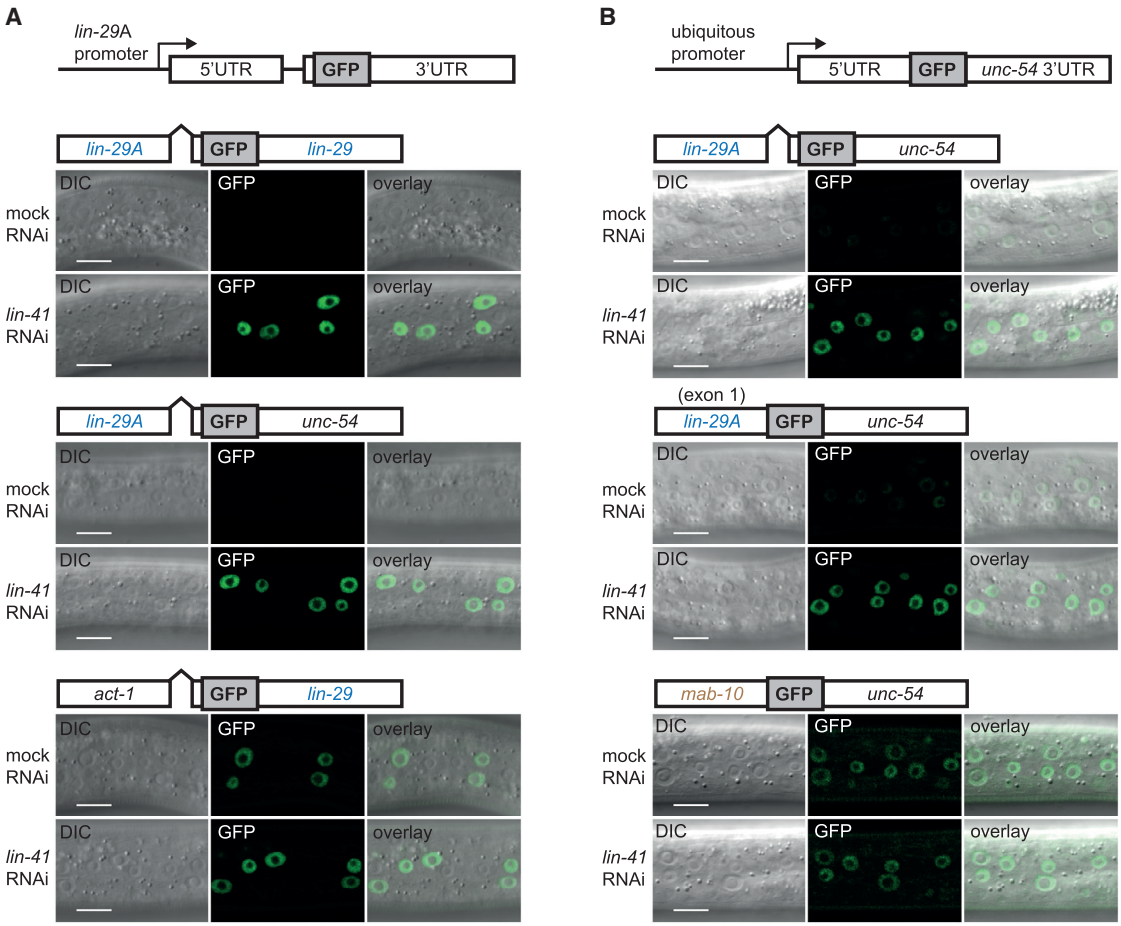


Figure 5. The 5′ UTR of *lin-29A* Mediates LIN41-Dependent Translational Repression

(A and B) Micrographs show early L3-stage animals, exposed to *lin-41* or mock RNAi, expressing nuclear-localized GFP reporters with the indicated 5′ and 3′ UTRs from (A) the *lin-29A* promoter or (B) the *dpy-30* promoter. See also Figure S4.

(Figure S5A). Using electrophoretic mobility shift assays (EMSAs), we found that the 194-nt-long 5′ UTR segment of *lin-29A* exon 1 was bound by recombinant LIN41, whereas a control fragment of the *unc-54* 3′ UTR was not (Figure 6A). Binding to the *lin-29A* 5′ UTR appears to involve multiple binding sites, since LIN41 bound to three different and partially overlapping ~100-nt fragments of it (Figures S5B and S5E). However, affinity of each fragment was reduced relative to the full-length fragment (Figure S5F), and, accordingly, none of the ~100-nt fragments sufficed for LIN41-mediated repression in vivo (Figure S6A).

We used six consecutive 200-nt fragments, overlapping by 50 nt, to test binding of LIN41 to the *mab-10* 3′ UTR. Whereas LIN41 failed to bind the first and the last two segments, it did bind each of three partially overlapping segments in the middle of the 3′ UTR (Figures 6B and 6C). As expected, the *mab-10* 5′ UTR was not bound by LIN41. The two non-overlapping *mab-10* 3′ UTR segments with clear EMSA shifts, parts 2 and 4, individually sufficed for repression of a GFP reporter when transplanted into the *unc-54* 3′ UTR (Figure 6D), confirming that there are at least two LIN41-binding sites on the *mab-10* 3′ UTR. When trying to delineate the minimal region needed for LIN41 binding to *mab-10* 3′ UTR parts 2 and 4, we found that LIN41 bound shorter ~100-nt RNA stretches poorly if at all (Figures S5C–S5E), similar to what we found with the *lin-29A* 5′ UTR. In conclusion, LIN41 has direct RNA-binding activity whose specificity in vitro reflects its target silencing specificity in vivo.

LIN41 Activity on *lin-29A* Differs from that on the Other Targets

To examine the silencing mechanism acting on reporter genes, we measured whole-worm *gfp* transcript levels for *lin-29A* promoter-containing constructs in early L3-stage worm lysates. As for endogenous *lin-29A* (Figure 4F), *lin-41*(*RNAi*) did not affect transcript levels of the reporter containing both *lin-29A* 5′ UTR and 3′ UTR (Figures 7A and S4C). We can exclude that this is due to use of whole-animal lysates, as they permit ready detection of LIN-29A repression by western blot (Figures 4E and S3D). Indeed, in early L3 stage, the reporter is silenced by LIN41 broadly across the epidermis, revealing LIN41 refractory expression in only a single cell, the anchor cell (data not shown). Hence, the absence

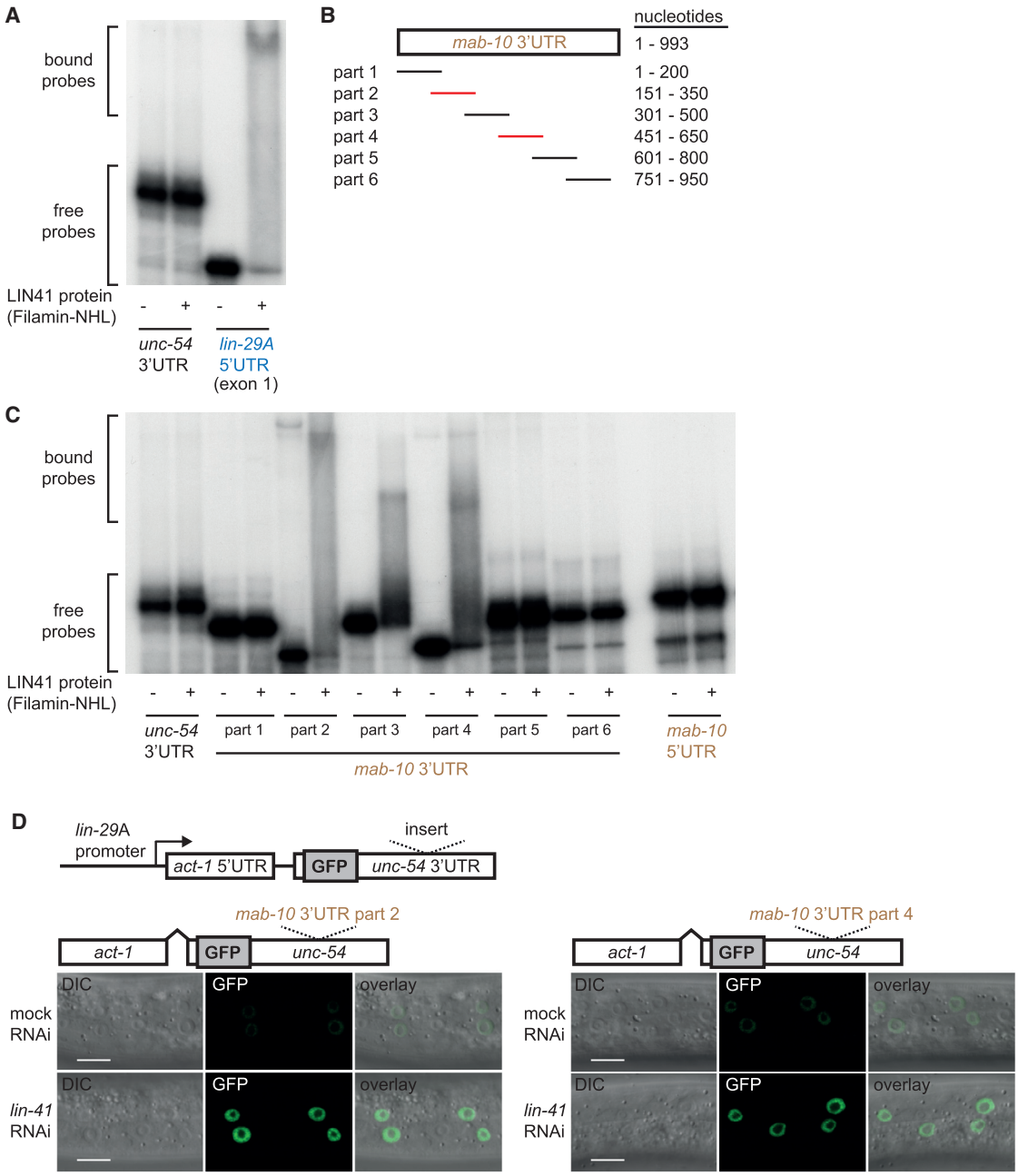


Figure 6. LIN41 Directly Binds to Its Target mRNAs

(A) Autoradiograph showing binding of LIN41 to the *lin-29A* 5′ UTR (194 nt). A 198-nt fragment from the *unc-54* 3′ UTR is used as a negative control.

(B) Schematic showing overlapping 200-nt-long radioactively labeled RNA gel-shift probes spanning the *mab-10* 3′ UTR, with those binding most efficiently to LIN41 in red.

(C) Autoradiograph showing binding of LIN41 to three 200-nt fragments (parts 2–4) from the *mab-10* 3′ UTR, but not to other *mab-10* 3′ UTR parts, the *mab-10* 5′ UTR (244 nt), or the control *unc-54* 3′ UTR fragment.

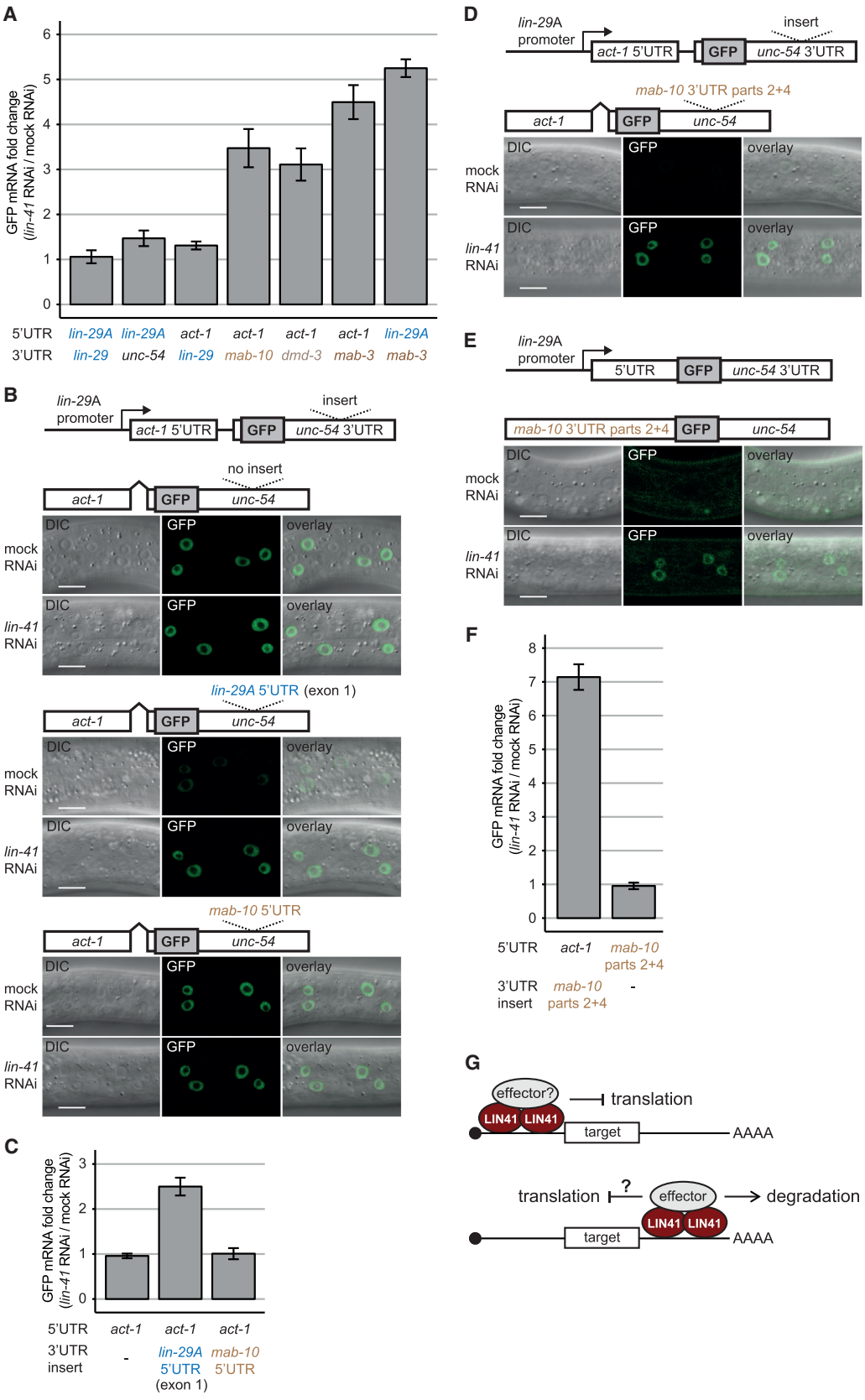
(D) Micrographs of early L3-stage animals, exposed to *lin-41* or mock RNAi, expressing nuclear-localized GFP reporters from the *lin-29A* promoter. The reporters contain the unregulated *act-1* 5′ UTR exon and the *unc-54* 3′ UTR, with either *mab-10* 3′ UTR part 2 or 4 as an insert.

For (A) and (C), the LIN41 protein used in the assays is N-terminally truncated and contains the Filamin and NHL domains.

See also Figures S5 and S6.

of detectable transcript degradation shows that the reporter recapitulates regulation of endogenous *lin-29A* through translational repression.

We envisioned three scenarios by which the *lin-29A* transcript might escape degradation by LIN41. First, LIN41 activity on this target might differ from that on other targets. Second, the *lin-29A*



(legend on next page)

3′ UTR and/or 5′ UTR might specifically protect the transcript from decay. Third, the epidermis as a major organ of *lin-29A* expression might lack the capacity to execute LIN41-mediated degradation altogether. To distinguish among these possibilities, we quantified transcript level changes for additional reporters (summarized in Figure S4F). First, we compared the reporter transgenes containing the *lin-29A* 5′ UTR and either the *lin-29* 3′ UTR or the *unc-54* 3′ UTR, both silenced at the GFP level (Figure 5A). Neither reporter revealed substantial LIN41-dependent changes in mRNA levels (Figure 7A), excluding a general stabilizing effect of the *lin-29* 3′ UTR. Next, we examined transcript levels of the reporter transgenes that shared the *lin-29A* promoter and *act-1* 5′ UTR but contained the 3′ UTRs of *lin-29*, *mab-10*, *dmd-3*, or *mab-3*, respectively (Figures 7A and S4B). Major LIN41-dependent decreases in transcript levels occurred with the latter three, excluding a lack of degradative activity in the epidermis. Finally, we tested a construct that contained the *lin-29A* 5′ UTR and the *mab-3* 3′ UTR, and we observed transcript degradation (Figures 7A and S4B). Thus, degradation prevailed over translational repression and the *lin-29A* 5′ UTR could not inhibit mRNA degradation. In summary, because there is LIN41-induced degradative activity in the epidermis and the reporters with *lin-29A* 5′ UTR or 3′ UTR are not inherently resistant to degradation, we conclude that, on *lin-29A*, the mode of repression by LIN41 truly differs from that on the other target genes.

Relocation of LIN41-Binding Sites Alters the Mode of Silencing

Given that *lin-29A* repression differed in both the location of the repressive element and the mechanism of action, we wondered if the former instructed the latter. Therefore, we examined if and how a reporter was repressed when the LIN41-binding site of the *lin-29A* 5′ UTR was placed in a 3′ UTR context. When transplanting the *lin-29A* 5′ UTR exon 1 into the *unc-54* 3′ UTR, we found that GFP expression from the resulting reporter transcript was indeed repressed (Figure 7B). Although the extent of repression seemed less than what was seen when the same element was present in the 5′ UTR, silencing was specific in that it depended on LIN41 and the inserted sequence. Strikingly, when examining the transcript levels of the reporter, we observed a substantial reduction in the presence of LIN41 (Figures 7C and S4D). Therefore, the same element that caused translational repression when present in a 5′ UTR specifically induced transcript degradation when occurring in a 3′ UTR.

To test if, conversely, transplantation of a 3′ UTR-binding site into a 5′ UTR sufficed to switch the mechanism of LIN41-mediated silencing from degradation to translational repression, we focused on the *mab-10* 3′ UTR. Although parts 2 and 4 each bound to LIN41 in vitro and, within a 3′ UTR, conferred reporter repression in vivo (Figures 6C and 6D), they were non-functional when individually used as a 5′ UTR, causing either no regulation or a general loss of translation (Figure S6B). However, when we combined them into a single fragment, the two parts sufficed for LIN41-dependent GFP repression, not only when placed in the heterologous *unc-54* 3′ UTR (Figure 7D) but also when utilized as a 5′ UTR (Figure 7E). In the ectopic 5′ UTR context, repression no longer relied on reporter gene degradation, indicating a switch to translational repression (Figures 7F and S4E). Hence, the two transplantation experiments demonstrate that the location of the LIN41-repressive element is a major determinant of the LIN41 mechanism of action (Figure 7G).

DISCUSSION

Previous work from in vitro cell culture revealed that LIN41 utilizes transcript degradation as a mechanism of action (Chang et al., 2012; Loedige et al., 2013; Mitschka et al., 2015). Here we confirm such activity for *mab-10*, *mab-3*, and *dmd-3* in vivo, but additionally we demonstrate a second and distinct activity, namely, translational repression. Other RBPs, most prominently the miRNA-guided Argonaute proteins, are capable of silencing target transcripts through translational repression and degradation (Jonas and Izaurralde, 2015). However, translational repression typically augments rather than replaces transcript degradation as a silencing mechanism (Jonas and Izaurralde, 2015), and target determinants that favor one mechanism over the other have remained elusive. By contrast, we find that LIN41 can silence a target purely by translational repression. Moreover, we identify the location of LIN41-binding sites on its targets as a key determinant for the choice between translational inhibition and mRNA degradation.

The position-dependent function of LIN41 was unexpected, but we propose that a systematic evaluation of RBP activities might reveal location-dependent choice of activity as a more common, presently underappreciated feature of RBPs. Indeed, although further instances of RBPs that execute translational repression and RNA degradation in a position-dependent manner remain to be uncovered, metazoan iron regulatory proteins (IRPs) and *Drosophila melanogaster* sex-lethal (SXL) are two additional examples of position-dependent RBP functions: IRPs inhibit translation initiation when binding to a 5′ UTR, but they protect mRNA from degradation when binding to a 3′ UTR (Kühn, 2015), and SXL utilizes two distinct mechanisms to repress translation of the *male-specific lethal* (*msl*)*-2* mRNA through its 5′ UTR and 3′ UTR, respectively (Beckmann et al., 2005).

Figure 7. The Binding Location of LIN41 Determines Its Mode of Repression

(A, C, and F) RT-qPCR analysis of *lin-29A* promoter-driven GFP reporter mRNA levels. Depicted are the fold changes of GFP mRNA levels (normalized by *act-1* mRNA levels) from early L3-stage worms grown on *lin-41* RNAi relative to those grown on mock RNAi bacteria. n = 3 biological replicates, data as mean ± SEM.

(B, D, and E) Micrographs of early L3-stage animals, exposed to *lin-41* or mock RNAi, expressing nuclear-localized GFP reporters from the *lin-29A* promoter. (B) The reporters contain the unregulated *act-1* 5′ UTR exon and the *unc-54* 3′ UTR, without insert or with either the *lin-29A* 5′ UTR exon or the *mab-10* 5′ UTR as an insert. (D and E) The reporters contain a 400-nt-long fragment with fused *mab-10* 3′ UTR parts 2 and 4, either (D) as an insert within the *unc-54* 3′ UTR or (E) as a 5′ UTR.

(G) Position-dependent modes of action are applied by LIN41 to repress its mRNA targets (see the Discussion for details).

See also Figures S6 and S7.

It seems likely that LIN41, like other RBPs, triggers mRNA degradation through recruitment of dedicated effector proteins, i.e., deadenylases and/or ribonucleases (Figure 7G). These may then be absent when LIN41 binds to 5′ UTRs, or additional modulators or steric constraints may prevent their activity. Translational silencing also may involve recruitment of a dedicated machinery. If present also when LIN41 binds targets at their 3′ UTRs, LIN41 may always impose a translational block, although we observe little or no robust translational repression of such targets (Figures S7A–S7C). In analogy to certain models of miRNA activity (Jonas and Izaurralde, 2015), however, translational repression might then be tightly coupled to degradation, making degradation the observable net outcome. Alternatively, translational repression may depend on context features that restrict it to the 5′ UTR. For instance, LIN41 on the 5′ UTR may operate according to a roadblock model, whereby its binding, alone or in a complex, impedes ribosomal recruitment or scanning.

Irrespective of the scenario that applies, our data provide insight into the mechanism of translational repression. The fact that LIN41 binding reduces RPF levels homogenously along the length of the *lin-29A* transcript (within exons 1–4 and 5–11, Figures 4A and S3A) argues against both a block in elongation and premature ribosome drop-off as possible mechanisms, and instead it implies regulation at the level of translation initiation. Increased translation of upstream open reading frames (uORFs), preventing (re-)initiation on the main ORF, is a well-established means of repressing translation initiation in a 5′ UTR-dependent manner (Hinnebusch, 2005; Medenbach et al., 2011). However, we do not expect uORFs to contribute to silencing of *lin-29A*. The *lin-29A* 5′ UTR lacks cognate AUG start codons, and no RPFs accumulate from this 5′ UTR when *lin-29A* translation is repressed (Figures S7D and S7E). (Although we note that, for reasons that remain to be established, RPFs do accumulate from this region when LIN41 is absent.) In addition, whereas repressive activities of uORFs are dependent on position relative to the main ORF AUG start codon, we found that insertion of 65 extra nucleotides from the *act-1* 5′ UTR into two different positions of the *lin-29A* 5′ UTR did not detectably perturb repression (Figure S6C). Finally, the in vitro mapping (Figures 6A–6C and S5) and in vivo reporter (Figures 6D, 7D, and S6A) experiments suggest that both target 3′ and 5′ UTRs contain multiple LIN41-binding sites distributed over >100 nt that are required for effective binding and efficient silencing.

The four in vivo targets that we have uncovered and validated, *mab-3*, *dmd-3*, *mab-10*, and *lin-29A*, are likely to be major physiological effectors of LIN41. All four are known heterochronic (temporal patterning) genes, as are *let-7* and *lin-41* (Rougvie and Moss, 2013), and *lin-41* interacts genetically with *lin-29* and *mab-10* in controlling skin development (Harris and Horvitz, 2011; Reinhart et al., 2000; Slack et al., 2000) and with *dmd-3* in male tail development (Mason et al., 2008). The data that we present here supply missing mechanistic links of the heterochronic pathway by demonstrating that all four genes are direct LIN41 targets and by revealing that LIN41 regulates only one *lin-29* isoform, *lin-29A*.

LIN-29A and MAB-10 are particularly interesting LIN41 targets: LIN-29A is an early growth response (EGR)-type transcription factor of the Krüppel family, whereas MAB-10, orthologous to mammalian NAB1/2 (NGFI-A-binding proteins 1 and 2), is its transcription cofactor (Harris and Horvitz, 2011). Like LIN41, EGR and NAB proteins regulate proliferation and/or terminal differentiation programs in various animals and cell types, as exemplified previously (Du et al., 2014; Laslo et al., 2006; Le et al., 2005; Min et al., 2008; Nguyen et al., 1993; Topilko et al., 1994). Most strikingly, *let-7*, LIN41, and EGR1 were all shown to affect reprogramming efficiency of mammalian epidermal fibroblasts into induced pluripotent stem cells in vitro (Worringer et al., 2014). Although a mechanism of LIN41-mediated repression of EGR1 has not been established, EGR1 mRNA co-immunoprecipitates with LIN41 from human embryonic stem cells (Worringer et al., 2014). Hence, LIN41 may regulate stem cell fates through an evolutionarily conserved effector pair, LIN-29A/EGR and MAB-10/NAB.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - C. elegans
- METHOD DETAILS
  - Ribosome Profiling and Total RNA Sequencing
  - Construction of GFP Reporters
  - Confocal Imaging and RNA Extraction
  - RNA Co-immunoprecipitation (RIP)
  - RT-qPCR
  - Western Blotting
  - Tagging of Endogenous *lin-29* by CRISPR-Cas9
  - Electrophoretic Mobility Shift Assay
- QUANTIFICATION AND STATISTICAL ANALYSES
  - Ribosome Profiling and RNA-Seq Data Analysis
  - RT-qPCR Analysis
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2016.12.010.

AUTHOR CONTRIBUTIONS

F.A. conceived the project; designed, performed, and computationally analyzed ribosome profiling and RNA-seq experiments; designed, performed, and analyzed confocal microscopy, RT-qPCR, and western blot experiments; created transgenic worm lines; and wrote the manuscript. P.K. purified the recombinant LIN41 protein fragment and performed and analyzed in vitro-binding experiments. H.B. performed RNA-IP experiments and created some transgenic worm lines. D.G. performed computational analysis of ribosome profiling and RNA-seq time courses. L.X. helped with cloning and worm injections. R.C. designed experiments. H.G. conceived the project, designed and analyzed experiments, and wrote the manuscript.

ACKNOWLEDGMENTS

We thank Monika Fasler and Magdalene Rausch for their help and support for this project. We are thankful to Iskra Katic for worm injections and comments on the manuscript, Jeremy Keusch for help in protein purification, and Giovanna Brancati for a reagent. We are grateful to Marc Buehler, Joerg Betschinger, Takashi Miki, Sarah Carl, Jun Liu, and Manuel de la Mata for discussions and comments on the manuscript. We thank Kirsten Jacobeit, Sophie Dessus-Babus, Stéphane Thiry, and Tim Roloff for help with library preparation and high-throughput sequencing and Laurent Gelman for help with confocal imaging. The work was partly supported by the Swiss National Science Foundation through grants 31003A_149402 (to R.C.), and grant 31003A_163447 and the NCCR RNA & Disease (to H.G.). Further funding for the research leading to these results was from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement 241985 (European Research Council "miRTurn") and the Novartis Research Foundation through the FMI (to H.G.). P.K. has received funding for the research leading to these results from an EMBO Fellowship (ALTF 95-2015) that was co-funded by the European Commission support for Marie Curie Actions (LTFCOFUND2013, GA-2013-609409) and FMI. Some strains were provided by the Caenorhabditis Genetics Center (CGC), which is funded by the NIH Office of Research Infrastructure Programs (P40 OD010440).

Received: July 26, 2016
Revised: November 11, 2016
Accepted: December 13, 2016
Published: January 19, 2017

SUPPORTING CITATIONS

The following references appear in the Supplemental Information: Wright et al. (2011); Merritt et al. (2008).

REFERENCES

Abrahante, J.E., Daul, A.L., Li, M., Volk, M.L., Tennessen, J.M., Miller, E.A., and Rougvie, A.E. (2003). The Caenorhabditis elegans hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs. Dev. Cell *4*, 625–637.

Aeschimann, F., Xiong, J., Arnold, A., Dieterich, C., and Großhans, H. (2015). Transcriptome-wide measurement of ribosomal occupancy by ribosome profiling. Methods *85*, 75–89.

Ahringer, J. (2006). Reverse genetics. WormBook, 1–43.

Ambros, V., and Horvitz, H.R. (1984). Heterochronic mutants of the nematode Caenorhabditis elegans. Science *226*, 409–416.

Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A.E. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. Cell *122*, 553–563.

Beckmann, K., Grskovic, M., Gebauer, F., and Hentze, M.W. (2005). A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in Drosophila. Cell *122*, 529–540.

Bettinger, J.C., Lee, K., and Rougvie, A.E. (1996). Stage-specific accumulation of the terminal differentiation factor LIN-29 during Caenorhabditis elegans development. Development *122*, 2517–2527.

Bettinger, J.C., Euling, S., and Rougvie, A.E. (1997). The terminal differentiation factor LIN-29 is required for proper vulval morphogenesis and egg laying in Caenorhabditis elegans. Development *124*, 4333–4342.

Brabin, C., and Woollard, A. (2012). Finding a niche for seam cells? Worm *1*, 107–111.

Chang, H.M., Martinez, N.J., Thornton, J.E., Hagan, J.P., Nguyen, K.D., and Gregory, R.I. (2012). Trim71 cooperates with microRNAs to repress Cdkn1a expression and promote embryonic stem cell proliferation. Nat. Commun. *3*, 923.

Chatterjee, S., and Großhans, H. (2009). Active turnover modulates mature microRNA activity in Caenorhabditis elegans. Nature *461*, 546–549.

Chen, J., Lai, F., and Niswander, L. (2012). The ubiquitin ligase mLin41 temporally promotes neural progenitor cell maintenance through FGF signaling. Genes Dev. *26*, 803–815.

Chiu, S.C., Chung, H.Y., Cho, D.Y., Chan, T.M., Liu, M.C., Huang, H.M., Li, T.Y., Lin, J.Y., Chou, P.C., Fu, R.H., et al. (2014). Therapeutic potential of microRNA let-7: tumor suppression or impeding normal stemness. Cell Transplant. *23*, 459–469.

Dickinson, D.J., Pani, A.M., Heppert, J.K., Higgins, C.D., and Goldstein, B. (2015). Streamlined genome engineering with a self-excising drug selection cassette. Genetics *200*, 1035–1049.

Ding, X.C., and Großhans, H. (2009). Repression of C. elegans microRNA targets at the initiation level of translation requires GW182 proteins. EMBO J. *28*, 213–222.

Du, N., Kwon, H., Li, P., West, E.E., Oh, J., Liao, W., Yu, Z., Ren, M., and Leonard, W.J. (2014). EGR2 is critical for peripheral naïve T-cell differentiation and the T-cell response to influenza. Proc. Natl. Acad. Sci. USA *111*, 16484–16489.

Ecsedi, M., and Großhans, H. (2013). LIN-41/TRIM71: emancipation of a miRNA target. Genes Dev. *27*, 581–589.

Ecsedi, M., Rausch, M., and Großhans, H. (2015). The let-7 microRNA directs vulval development through a single target. Dev. Cell *32*, 335–344.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. *30*, 207–210.

Evans, T.C. (2006). Transformation and microinjection. WormBook, 1–15.

Fraser, A.G., Kamath, R.S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. Nature *408*, 325–330.

Frøkjaer-Jensen, C., Davis, M.W., Hopkins, C.E., Newman, B.J., Thummel, J.M., Olesen, S.P., Grunnet, M., and Jorgensen, E.M. (2008). Single-copy insertion of transgenes in Caenorhabditis elegans. Nat. Genet. *40*, 1375–1383.

Frøkjær-Jensen, C., Davis, M.W., Ailion, M., and Jorgensen, E.M. (2012). Improved Mos1-mediated transgenesis in C. elegans. Nat. Methods *9*, 117–118.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Großhans, H., Johnson, T., Reinert, K.L., Gerstein, M., and Slack, F.J. (2005). The temporal patterning microRNA let-7 regulates several transcription factors at the larval to adult transition in C. elegans. Dev. Cell *8*, 321–330.

Harris, D.T., and Horvitz, H.R. (2011). MAB-10/NAB acts with LIN-29/EGR to regulate terminal differentiation and the transition from larva to adult in C. elegans. Development *138*, 4051–4062.

Hendriks, G.J., Gaidatzis, D., Aeschimann, F., and Großhans, H. (2014). Extensive oscillatory gene expression during C. elegans larval development. Mol. Cell *53*, 380–392.

Hinnebusch, A.G. (2005). Translational regulation of GCN4 and the general amino acid control of yeast. Annu. Rev. Microbiol. *59*, 407–450.

Ikeda, K., and Inoue, S. (2012). TRIM proteins as RING finger E3 ubiquitin ligases. Adv. Exp. Med. Biol. *770*, 27–37.

Jonas, S., and Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. Nat. Rev. Genet. *16*, 421–433.

Joshi, P.M., Riddle, M.R., Djabrayan, N.J., and Rothman, J.H. (2010). Caenorhabditis elegans as a model for stem cell biology. Dev. Dyn. *239*, 1539–1554.

Katic, I., Xu, L., and Ciosk, R. (2015). CRISPR/Cas9 genome editing in Caenorhabditis elegans: evaluation of templates for homology-mediated repair and knock-ins by homology-independent DNA repair. G3 (Bethesda) *5*, 1649–1656.

Kühn, L.C. (2015). Iron regulatory proteins and their role in controlling iron metabolism. Metallomics *7*, 232–243.

Kwon, S.C., Yi, H., Eichelbaum, K., Föhr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., and Kim, V.N. (2013). The RNA-binding protein repertoire of embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1122–1130.

Laslo, P., Spooner, C.J., Warmflash, A., Lancki, D.W., Lee, H.J., Sciammas, R., Gantner, B.N., Dinner, A.R., and Singh, H. (2006). Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. Cell *126*, 755–766.

Le, N., Nagarajan, R., Wang, J.Y., Svaren, J., LaPash, C., Araki, T., Schmidt, R.E., and Milbrandt, J. (2005). Nab proteins are essential for peripheral nervous system myelination. Nat. Neurosci. *8*, 932–940.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402–408.

Loedige, I., Gaidatzis, D., Sack, R., Meister, G., and Filipowicz, W. (2013). The mammalian TRIM-NHL protein TRIM71/LIN-41 is a repressor of mRNA function. Nucleic Acids Res. *41*, 518–532.

Loedige, I., Stotz, M., Qamar, S., Kramer, K., Hennig, J., Schubert, T., Löffler, P., Längst, G., Merkl, R., Urlaub, H., and Meister, G. (2014). The NHL domain of BRAT is an RNA-binding domain that directly contacts the hunchback mRNA for regulation. Genes Dev. *28*, 749–764.

Loedige, I., Jakob, L., Treiber, T., Ray, D., Stotz, M., Treiber, N., Hennig, J., Cook, K.B., Morris, Q., Hughes, T.R., et al. (2015). The crystal structure of the NHL domain in complex with RNA reveals the molecular basis of Drosophila brain-tumor-mediated gene regulation. Cell Rep. *13*, 1206–1220.

Löer, B., Bauer, R., Bornheim, R., Grell, J., Kremmer, E., Kolanus, W., and Hoch, M. (2008). The NHL-domain protein Wech is crucial for the integrin-cytoskeleton link. Nat. Cell Biol. *10*, 422–428.

Mason, D.A., Rabinowitz, J.S., and Portman, D.S. (2008). dmd-3, a doublesex-related gene regulated by tra-1, governs sex-specific morphogenesis in C. elegans. Development *135*, 2373–2382.

Medenbach, J., Seiler, M., and Hentze, M.W. (2011). Translational control via protein-regulated upstream open reading frames. Cell *145*, 902–913.

Merritt, C., Rasoloson, D., Ko, D., and Seydoux, G. (2008). 3′ UTRs are the primary regulators of gene expression in the C. elegans germline. Curr Biol *18*, 1476–1482.

Min, I.M., Pietramaggiori, G., Kim, F.S., Passegué, E., Stevenson, K.E., and Wagers, A.J. (2008). The transcription factor EGR1 controls both the proliferation and localization of hematopoietic stem cells. Cell Stem Cell *2*, 380–391.

Mitschka, S., Ulas, T., Goller, T., Schneider, K., Egert, A., Mertens, J., Brüstle, O., Schorle, H., Beyer, M., Klee, K., et al. (2015). Co-existence of intact stemness and priming of neural differentiation programs in mES cells lacking Trim71. Sci. Rep. *5*, 11126.

Nguyen, H.Q., Hoffman-Liebermann, B., and Liebermann, D.A. (1993). The zinc finger transcription factor Egr-1 is essential for and restricts differentiation along the macrophage lineage. Cell *72*, 197–209.

Rehfeld, F., Rohde, A.M., Nguyen, D.T., and Wulczyn, F.G. (2015). Lin28 and let-7: ancient milestones on the road from pluripotency to neurogenesis. Cell Tissue Res. *359*, 145–160.

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature *403*, 901–906.

Rougvie, A.E., and Ambros, V. (1995). The heterochronic gene lin-29 encodes a zinc finger protein that controls a terminal differentiation event in Caenorhabditis elegans. Development *121*, 2491–2500.

Rougvie, A.E., and Moss, E.G. (2013). Developmental transitions in C. elegans larval stages. Curr. Top. Dev. Biol. *105*, 153–180.

Rüegger, S., Miki, T.S., Hess, D., and Großhans, H. (2015). The ribonucleotidyl transferase USIP-1 acts with SART3 to promote U6 snRNA recycling. Nucleic Acids Res. *43*, 3344–3357.

Rybak, A., Fuchs, H., Hadian, K., Smirnova, L., Wulczyn, E.A., Michel, G., Nitsch, R., Krappmann, D., and Wulczyn, F.G. (2009). The let-7 target gene mouse lin-41 is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. Nat. Cell Biol. *11*, 1411–1420.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Slack, F.J., Basson, M., Liu, Z., Ambros, V., Horvitz, H.R., and Ruvkun, G. (2000). The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. Mol. Cell *5*, 659–669.

Spike, C.A., Coetzee, D., Eichten, C., Wang, X., Hansen, D., and Greenstein, D. (2014a). 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*, 1535–1558.

Spike, C.A., Coetzee, D., Nishi, Y., Guven-Ozkan, T., Oldenbroek, M., Yamamoto, I., Lin, R., and Greenstein, D. (2014b). Translational control of the oogenic program by components of OMA ribonucleoprotein particles in Caenorhabditis elegans. Genetics *198*, 1513–1533.

Tocchini, C., and Ciosk, R. (2015). TRIM-NHL proteins in development and disease. Semin. Cell Dev. Biol. *47-48*, 52–59.

Tocchini, C., Keusch, J.J., Miller, S.B., Finger, S., Gut, H., Stadler, M.B., and Ciosk, R. (2014). The TRIM-NHL protein LIN-41 controls the onset of developmental plasticity in Caenorhabditis elegans. PLoS Genet. *10*, e1004533.

Topilko, P., Schneider-Maunoury, S., Levi, G., Baron-Van Evercooren, A., Chennoufi, A.B., Seitanidou, T., Babinet, C., and Charnay, P. (1994). Krox-20 controls myelination in the peripheral nervous system. Nature *371*, 796–799.

Vella, M.C., Choi, E.Y., Lin, S.Y., Reinert, K., and Slack, F.J. (2004). The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. Genes Dev. *18*, 132–137.

Worringer, K.A., Rand, T.A., Hayashi, Y., Sami, S., Takahashi, K., Tanabe, K., Narita, M., Srivastava, D., and Yamanaka, S. (2014). The let-7/LIN-41 pathway regulates reprogramming to human induced pluripotent stem cells by controlling expression of prodifferentiation genes. Cell Stem Cell *14*, 40–52.

Wright, J.E., and Ciosk, R. (2013). RNA-based regulation of pluripotency. Trends Genet. *29*, 99–107.

Wright, J.E., Gaidatzis, D., Senften, M., Farley, B.M., Westhof, E., Ryder, S.P., and Ciosk, R. (2011). A quantitative RNA code for mRNA target selection by the germline fate determinant GLD-1. EMBO J *30*, 533–545.

Ye, J., and Blelloch, R. (2014). Regulation of pluripotency by RNA binding proteins. Cell Stem Cell *15*, 271–280.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Monoclonal mouse anti-FLAG M2-Peroxidase (HRP)	Sigma-Aldrich	Cat#A8592, RRID: AB_439702
Monoclonal mouse anti-Actin clone C4	Millipore	Cat#MAB1501, RRID: AB_2223041
Horseradish peroxidase-conjugated secondary antibody	GE Healthcare	Cat#NXA931, RRID: AB_772209
Chemicals, Peptides, and Recombinant Proteins		
StrepTag-TEV-LIN41_Filamin_NHL_Domain (*C. elegans*)	This study	N/A
Critical Commercial Assays		
ScriptSeq v2 RNA-Seq library preparation kit	Epicenter	Cat#SSV21124
Ribo-Zero rRNA Removal Kit	Epicenter	Cat#MRZH11124
TruSeq Small RNA Sample Preparation Kit	Illumina	Cat#RS-200-0012
Deposited Data		
Ribosome profiling and RNA sequencing data for wild-type run in parallel with *let-7(n2853)*	Hendriks et al., 2014	GEO: GSE52864 (GSM1277189-GSM1277198) and GEO GSE52905
Ribosome profiling and RNA sequencing data	This study	GEO: GSE80159
Raw microscopy image data related to main figures	This study	doi: 10.17632/wkcr5gb4t5.1 (Mendeley)
Experimental Models: Organisms/Strains		
C. elegans lines are listed in Table S4	N/A	N/A
Recombinant DNA		
Plasmids are listed in Table S5	N/A	N/A
Primers are listed in Tables S6 (qPCR) and S7 (EMSA probe generation)	N/A	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Helge Großhans, Friedrich Miescher Institute for Biomedical Research, helge.grosshans@fmi.ch, will respond to request and provide reagents and information. Published research reagents from the FMI are shared with the academic community under a Material Transfer Agreement (MTA) having terms and conditions corresponding to those of the UBMTA (Uniform Biological Material Transfer Agreement).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

C. elegans

The worm strains used in this study are listed in Table S4. The wild-type strain was Bristol N2. To synchronize worms, arrested L1 stage larvae were obtained by extracting embryos from gravid adults using a bleaching solution (30% (v/v) sodium hypochlorite (5% chlorine) reagent (Thermo Fisher Scientific; 419550010), 750 mM KOH). *let-7(n2853)* embryos were extracted from mothers grown at permissive temperature (15°C). Synchronized arrested L1 larvae (L1s) were obtained by hatching overnight in the absence of food, at room temperature in M9 buffer (42 mM Na_2HPO_4, 22 mM KH_2PO_4, 86 mM NaCl, 1 mM $MgSO_4$), plated on food and incubated at 25°C for the desired time (termed hours of development in the respective experiments). For ribosome profiling or RNA-IP experiments, L1s were plated on enriched peptone plates with *Escherichia coli* NA22 bacteria (Evans, 2006). For RNAi experiments, L1s were plated on RNAi-inducing NGM agar plates with *Escherichia coli* HT115 bacteria containing plasmids targeting the gene of interest (Ahringer, 2006).

METHOD DETAILS

Ribosome Profiling and Total RNA Sequencing

Ribosome profiling time course experiments were performed according to our detailed published protocol (Aeschimann et al., 2015). In brief, synchronized worms, grown on enriched peptone plates with NA22 bacteria, were harvested every two hours from 18 hr to 36 or 38 hr, respectively, of development at 25°C. Between 200,000 worms (earliest time points) and 100,000 worms (latest time points) were collected. Worm lysates of 11 absorbance units at 260 nm were prepared in a total volume of 385 μl and digested with 2 μl of RNaseI (100 Units/μl, Life Technologies; AM2295) for 1 hr at 23°C. Monosomes were purified using linear sucrose density gradients for the first time course experiment (wild-type and *let-7(n2853)* animals) and using size-exclusion chromatography for the second time course experiment (wild-type, *lin-41(xe11)* and *lin-41(xe11); let-7(n2853)* animals), as described in (Aeschimann et al., 2015). RPFs were obtained by separation of the monosomal RNA on Novex 15% (w/v) Polyacrylamide TBE-Urea Gels (Life Technologies; EC6885BOX) and extraction of 28-30 nt long RNAs. Library preparation was performed according to the TruSeq Small RNA Sample Preparation Kit (Illumina; RS-200-0012), adjusted as described in (Aeschimann et al., 2015). For total RNA sequencing, a sample of the input RNA (before the RNase digest) was extracted using Tri Reagent (Molecular Research Center; TR 118) according to the manufacturer's recommendations. To obtain ribosomal RNA (rRNA)-depleted total RNA, a DNase-treatment was performed with the RNase-Free DNase Set (QIAGEN; 79254) and the RNeasy MiniKit (QIAGEN; 74104), before using the Ribo-Zero rRNA Removal Kit (Epicenter; MRZH11124) to remove rRNA. Libraries were prepared with the ScriptSeq v2 RNA-Seq library preparation kit (Epicenter; SSV21124) and, like RPF libraries, sequenced on an Illumina HiSeq2000 machine. The data of the first wild-type animal time course, collected for the ribosome profiling experiment with wild-type and *let-7(n2853)* worms, was first published elsewhere (Hendriks et al., 2014), GEO: GSE52864 (GSM1277189-GSM1277198) and GEO: GSE52905. Data analysis is described in the relevant section below.

Construction of GFP Reporters

All reporters were constructed using the MultiSite Gateway Technology (Thermo Fisher Scientific) and the destination vector pCFJ150 (Frøkjaer-Jensen et al., 2008). First, promoters, 5′UTRs and 3′UTRs were amplified from *C. elegans* genomic DNA or ordered as gBlocks® Gene Fragments (Integrated DNA Technologies), before inserting them into Entry clones using the Gateway BP Clonase II Enzyme mix (Thermo Fisher Scientific; 11789020) or Gibson assembly (Gibson et al., 2009). PCR primer or gBlock sequences, cloning techniques and resulting Entry plasmids are listed in Table S5. Second, three entry plasmids were recombined with the pCFJ150 vector backbone (Gateway LR Clonase II Enzyme mix, Thermo Fisher Scientific; 11791020) to a plasmid with promoter, 5′UTR, GFP(PEST)-H2B coding sequence and 3′UTR. Third, transgenic worms were obtained by single-copy integration into the *ttTi5605* locus on chromosome II, following the published protocol for injection with low DNA concentration (Frøkjær-Jensen et al., 2012).

Confocal Imaging and RNA Extraction

Before subjecting worms to RNA extraction or confocal imaging, they were grown for 20 hr at 25°C on RNAi-inducing plates with HT115 bacteria, either containing the insert-less L4440 parental RNAi vector (denoted "mock RNAi") or an RNAi vector with an insert targeting *lin-41* (Fraser et al., 2000). For RNA extraction, worms were harvested and washed with M9 buffer and frozen in 1 mL of Tri Reagent at −80°C until further use. RNA was extracted according to the Tri Reagent manufacturer's protocol, following lysis of worms with five repeats of freeze and thaw cycles using liquid nitrogen and a heating block at 42°C. For confocal imaging, worms were mounted on a 2% (w/v) agarose pad with a drop of 10 mM levamisole solution, and imaged on a Zeiss LSM 700 confocal microscope driven by Zen 2012 Software. Before acquiring images of representative worms, the GFP signals for at least 10 worms were observed to verify that they were comparable among different worms in each worm line and for each condition. A second independent integrant line was obtained for each construct and examined to confirm results. Fluorescent and Differential Interference Contrast (DIC) images were acquired with a 40x/1.3 oil immersion objective (1024x1024 pixels, pixel size 156nm). Selections of representative regions and processing of images was performed with Fiji (Schindelin et al., 2012). Identical worm lines grown on mock or *lin-41* RNAi bacteria were imaged and processed with identical settings. In LIN41-depleted worms of the L3 stage used for imaging, nuclear GFP reporters driven from the *dpy-30* promoter accumulated in seam cell nuclei as well as in the surrounding hypodermal (hyp7) nuclei, while those driven from the *lin-29A* promoter only accumulated in hyp7 nuclei. The latter expression pattern was also observed for endogenously tagged LIN-29 during L3 stage. LIN-29 in wild-type worms accumulates in seam cells starting only in the L4 stage (Bettinger et al., 1996; data not shown).

RNA Co-immunoprecipitation (RIP)

RIP was performed with non-transgenic wild-type worms, wild-type worms expressing *flag::gfp::sart-3* (Rüegger et al., 2015) and *lin-41(n2914)* mutant worms expressing *flag::gfp::lin-41*. The transgene for expression of *flag::gfp::lin-41* was cloned using MultiSite Gateway Technology and single-copy integrated on chromosome II, as described for the construction of GFP reporters (primers and Entry plasmids are listed in Table S5). The transgenic line was outcrossed four times to the wild-type strain before crossing it into the *lin-41(n2914)* mutant background. Worms with transgenic expression of FLAG::GFP::LIN41 in the *lin-41(n2914)* mutant background were superficially wild-type, and did not show any of the *lin-41(n2914)* mutant phenotypes (sterility, lethality, dumpiness).

Worms were harvested as semi-synchronous L3/L4 stage populations, obtained by bleaching of gravid adults, followed by directly plating the extracted embryos on enriched peptone plates with NA22 bacteria and incubating them for approximately 30 hr at 25°C. Worm pellets of about 1 mL were lysed in extraction buffer (50 mM HEPES/KOH (pH 7.4 at 4°C), 150 mM KCl, 5 mM $MgCl_2$, 0.1% (v/v) Triton X-100, 10% (w/v) glycerol, 1 mM PMSF, 7 mg/ml cOmplete Protease Inhibitor Tablets (EDTA-free, Roche; 11873580001), 200 U/ml RNase inhibitor (e.g., SUPERase In RNase Inhibitor, Life Technologies; AM2696)), with mortar and pestle in the presence of liquid nitrogen (see also (Aeschimann et al., 2015)). Lysates were cleared by centrifugation at 12,000 g for 10 min at 4°C. Anti-FLAG IPs were performed by incubating 3 mg total protein with 30 μl of anti-FLAG M2 magnetic beads (Sigma–Aldrich; M8823) for 3 hr at 4°C on a rotating wheel. Beads were washed five times for five minutes in extraction buffer without protease and RNase inhibitors, before extracting the bound RNA by directly adding Tri Reagent to the beads. For each condition, five IPs were performed in parallel to increase the amount of recovered RNA. A sample of input RNA was extracted from an aliquot of remaining input lysate using Tri Reagent. Reverse transcription (see below) was performed with 900 ng of input RNA and with 50% of the IP RNA, respectively. (IP RNA was not quantified due to low amounts.) After RT-qPCR analysis, a relative enrichment ("re") in IP versus input was calculated for each measured mRNA, separately for LIN41 IP, SART-3 IP and wild-type IP: re = 2^(CT (input) - CT (IP)). Fold enrichments compared to mock IP were then calculated by dividing LIN41 and SART-3 IP "re" values by the wild-type IP "re" values.

RT-qPCR

Reverse transcription was performed with the ImpromII Reverse Transcription System (Promega; A3800), according to the manufacturer's protocol, with 900 ng RNA (except for RNA from RIPs) and random primers (Promega; C1181). Using SYBR Green PCR Master Mix (Thermo Fisher Scientific; 4309155), qPCR was performed on a StepOnePlus Real-Time PCR System (Thermo Fisher Scientific) with the primers listed in Table S6. For comparing mRNA levels of the GFP reporters or of GFP-3xFLAG-tagged *lin-29*, grown on *lin-41* or mock RNAi bacteria, GFP mRNA fold changes were calculated with the $2^{-\Delta\Delta CT}$ Method (Livak and Schmittgen, 2001), using *act-1* as an internal control mRNA and the mock RNAi condition as calibrator.

Western Blotting

Worms were grown for 20 hr at 25°C on RNAi-inducing plates, as described above for confocal imaging and RNA extraction. Lysates were made by boiling (5 min, 95°C) and sonication in SDS lysis buffer (63 mM Tris-HCl (pH 6.8), 5 mM DTT, 2% SDS, 5% sucrose) and cleared by centrifugation, before separating proteins by SDS-PAGE (loading: 50 μg protein extract per well) and transferring them to PVDF membranes by semi-dry blotting. The following antibodies were used: Monoclonal mouse anti-FLAG M2-Peroxidase (HRP) (Sigma-Aldrich; A8592, dilution: 1:1,000). Monoclonal mouse anti-Actin clone C4 (Millipore; MAB1501, dilution 1:10,000). A horseradish peroxidase-conjugated secondary antibody (NXA931), ECL Western Blotting Detection Reagents and an ImageQuant LAS 4000 chemiluminescence imager (all from GE Healthcare) were used for signal generation and detection, respectively.

Tagging of Endogenous *lin-29* by CRISPR-Cas9

Wild-type worms were injected with a mix of 50 ng/μl pIK155, 100 ng/μl of pIK198 with a cloned sgRNA (atattatttatcagtgattg), 2.5 ng/μl pCFJ90, 5 ng/μl pCFJ104 and 10 ng/μl pDD282 with cloned homology arms, as described in (Dickinson et al., 2015; Katic et al., 2015). The plasmid for homologous recombination was prepared by restriction digest of pDD282 with ClaI and SpeI, followed by a 3-fragment Gibson assembly reaction (Gibson et al., 2009) with two gBlocks® Gene Fragments (Integrated DNA Technologies), as described in Table S5. Recombinants were isolated according to the protocol by Dickinson et al. (Dickinson et al., 2015), verified by DNA sequencing and outcrossed three times. Two independent worm lines were obtained and characterized.

Electrophoretic Mobility Shift Assay

Radioactively labeled probes for EMSAs were transcribed from PCR products with T3 RNA polymerase. Templates for probe synthesis were generated by PCR with an extended phage T3 RNA polymerase promoter (AATTAACCCTCACTAAAGGGAGAA) appended to the 5′end of the 5′primer, and gel-purified (primers are listed in Table S7). Labeled probes were transcribed in 3 μl reactions containing 0.5 μl template, 1.5 μl $^{\alpha}$P32 UTP (3 μM) (Hartmann Analytic), 0.6 μl 5x transcription buffer (Promega), 0.4 μl T3 polymerase (Promega), 0.2 μl RNasin (Promega), 2.5 mM rATP, rGTP and rCTP, and 0.025 mM rUTP (Roche) at 37°C for 3 hr. The reaction was stopped by adding 40 μl Tris-EDTA buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA) containing 30% glycerol and ~0.01% Bromophenol Blue. The C-terminal part of LIN41 containing the Filamin and NHL domains was produced as a Strep-tagged fusion protein in Sf9 insect cells using the baculovirus expression system (Invitrogen). The recombinant protein was affinity-purified by the Strep tag using Macroprep Strep-Tactin beads (IBA) according to the manufacturer's instruction. Protein was concentrated using ultra centrifugal filters (Amicon) and stored in aliquots at −80°C.

1 μl of 5 μM protein was pre-incubated with 4 μl of 2x gel-shift buffer (20 mM HEPES pH 8, 100 mM KCl, 200 mM NaCl, 0.2 mM EDTA, 20 mM DTT, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.2 mM $ZnSO_4$, 60% glycerol, 500 μg/ml heparin, 50 μg/ml *E. coli* tRNA). The reaction was made up to 7 μl with sterile water, incubated for 10 min at room temperature, following which 1 μl of RNA probe (~2 nM, ~10^5 cpm) was added. The reaction was incubated for 20 min and loaded onto the gel, electrophoresed at 25 mA, dried, and auto-radiographed. The cathode buffer was 50 mM glycine, 6 mM TRIS, 0.2 mM EDTA, pH 8), the stacking gel was 25 mM TRIS pH 6.8, 3% 19:1 acrylamide:bisacrylamide, the resolving gel was 0.5 X TBE (45 mM Tris-borate pH 8.3, 1 mM EDTA), 6% 19:1 acrylamide:bisacrylamide, and the anode buffer was 0.5X TBE. For competition gel-shift assays, increasing amounts of cold (unlabeled) competitor RNA

was added to the gel-shift reaction described above. 12 nM, 60 nM or 300 nM of cold RNA was incubated with LIN-41_Fil_NHL protein in 1x gel-shift buffer for 20 min, following which 1 μl of RNA probe (~2 nM, ~10^5 cpm) was added. The reaction was further incubated for 20 min and loaded onto the gel.

QUANTIFICATION AND STATISTICAL ANALYSES

Ribosome Profiling and RNA-Seq Data Analysis

RPF as well as RNA-seq data were processed and normalized as previously described (Hendriks et al., 2014), with separate normalizations for each batch of time course experiments. In all figures, expression levels are depicted as normalized $\log_2$ read counts. Before $\log_2$ transformation, a pseudocount of 8 was added to minimize large differences in expression caused by genes with a low number of read counts. Zero read counts thus result in a $\log_2$ read count of $\log_2(8) = 3$. The comparison of gene expression changes in the different mutants was performed with normalized $\log_2$ RPF counts, averaged for the five time points at 28, 30, 32, 34, and 36 hr (*lin-41* repression plateau). Consistently dysregulated genes at the RPF level in *let-7(n2853)* and *lin-41(xe11)* as compared to wild-type animals were identified based on the scatterplot depicted in Figure 1D. To compensate for the lower extent of *lin-41* dysregulation relative to the *let-7(n2853)* situation, we amplified the $\log_2$ fold changes in the *lin-41(xe11)* mutant by a factor of 1.5 for further processing. We first selected the up- and downregulated genes in the two mutants by requiring an average (adjusted) $\log_2$ fold change of at least 0.85. Second, to remove genes that were selected because they were mainly dysregulated in only one of the two mutants, we excluded genes with more than ± 30 degree angular deviation from the diagonal. To examine temporal changes of the selected up- and downregulated genes, the RPF read profiles were visually examined to determine the earliest developmental time point in which each gene was clearly and consistently dysregulated in all of the three mutant animals (*let-7(n2853)*, *lin-41(xe11)* and *lin-41(xe11); let-7(n2853)*, Table S3). To analyze RPF and RNA-seq data for the different exons of the *lin-29* gene, reads were counted in each of the 11 exons of the *lin-29A* (W03C9.4a) isoform separately (Rougvie and Ambros, 1995). Reads were normalized as above, with each exon treated as if it were a separate gene. In order to analyze the expression pattern of the exons specific to *lin-29A* versus all other exons (Figure 4B), the normalized reads for exons 1-4 and for exons 5-11 were summed up before addition of a pseudocount of 8 and $\log_2$ transformation. In order to analyze the expression changes for each separate exon during the *lin-41* repression plateau (Figure 4A), normalized RPF and mRNA read counts were summed up for the five time points at 28, 30, 32, 34 and 36 hr. Next, the read sums were $\log_2$ transformed, after addition of a pseudocount of 2. (Here, we chose a lower pseudocount to be able to visualize fold changes even with low read numbers per exon.) The fold downregulation for RPF and mRNA reads, respectively, was then quantified as differences in $\log_2$ read sums between wild-type and mutant. The non-coding exon 1 (pure 5′UTR sequence, no ATG start codon present) was included in this analysis, because we observed RPFs mapping to it, although at lower numbers than to coding exons. RPF reads in 5′UTRs have been observed in many studies, including our own metagene analysis (Aeschimann et al., 2015). While the nature of the reads mapping to *lin-29* exon 1 is unclear to us, they only accumulate in the absence of LIN41.

RT-qPCR Analysis

All RT-qPCR data are depicted as the mean of n = 3 or n = 4 biological replicates, with error bars representing ± s.e.m. In this study, a biological replicate is defined as an independently grown worm population, before extraction of embryos from gravid adults to obtain synchronized or semi-synchronized populations of worms (see EXPERIMENTAL MODEL AND SUBJECT DETAILS). The exact values of n are indicated in the figure legends.

For all reporter experiments, GFP or *lin-41* mRNA fold changes are calculated using the $2^{-\Delta\Delta CT}$ Method (Livak and Schmittgen, 2001), with *act-1* as an internal control mRNA and mock RNAi condition as calibrator. For each measured mRNA in each RNA-IP, a relative enrichment (“re”) in IP versus input is calculated: re = 2^(CT (input) - CT (IP)). The depicted fold enrichments compared to mock IP are calculated by dividing LIN41 and SART-3 IP “re” values by the wild-type IP “re” values. See also the relevant sections in the METHOD DETAILS.

DATA AND SOFTWARE AVAILABILITY

All ribosome profiling and RNA-sequencing data generated in this study have been deposited in the NCBI Gene Expression Omnibus (Edgar et al., 2002) under GEO: GSE80159. The wild-type time course data used for the comparison to *let-7(n2853)* have been previously deposited under accession numbers GEO: GSE52864 (GSM1277189-GSM1277198) and GEO: GSE52905. Raw microscopy image data for all main figures have been deposited at Mendeley, doi: 10.17632/wkcr5gb4t5.1.

Molecular Cell, Volume *65*

Supplemental Information

LIN41 Post-transcriptionally Silences mRNAs by Two Distinct and Position-Dependent Mechanisms

Florian Aeschimann, Pooja Kumari, Hrishikesh Bartake, Dimos Gaidatzis, Lan Xu, Rafal Ciosk, and Helge Großhans

SUPPLEMENTAL FIGURES

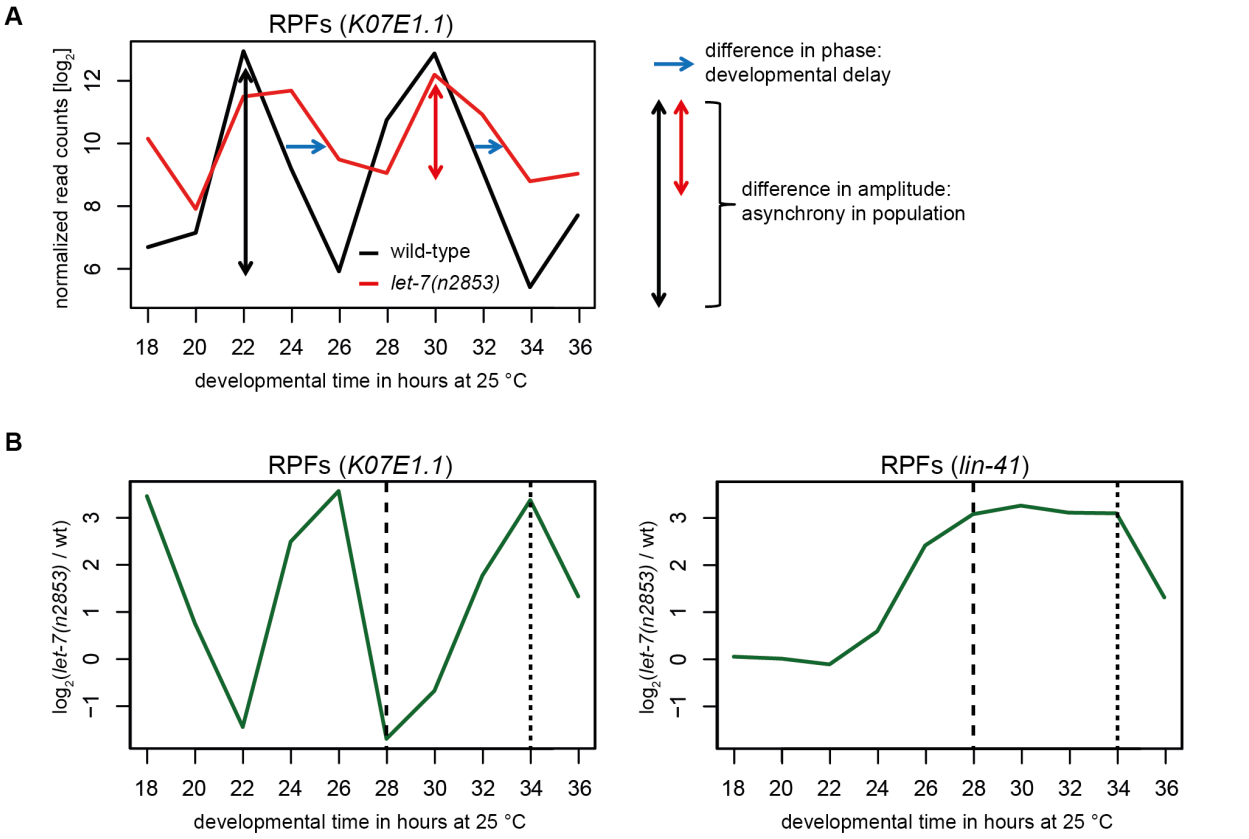


Figure S1. Related to Figure 1. Genes oscillating during development in differential gene expression analysis
(A) Expression of a highly oscillating gene (*K07E1.1*) over development at the level of ribosome-protected fragments (RPFs). Amplitude and phase of the oscillation are different between the two worm strains, as *let-7(n2853)* worms develop slightly more slowly and less synchronously than wild-type worms.
(B) Expression fold changes in the *let-7(n2853)* mutant animals compared to wild-type animals, plotted over developmental time for the highly oscillating gene *K07E1.1* and for the known *let-7* target *lin-41*. Analysis of fold changes in gene expression at a specific time point could result in misinterpretation of the data, as highly oscillating genes would be selected as differentially expressed. In the example, both *K07E1.1* and *lin-41* would be considered as about 8-fold higher ($\log_2(8) = 3$) in expression in *let-7(n2853)* mutant animals compared to wild-type animals if analyzed only at 34 h of development (dashed lines with short dashes). If analyzed at 28 h of development (dashed lines with long dashes), *lin-41* would still be considered 8-fold higher in expression, but *K07E1.1* would be considered as almost 4-fold down-regulated in *let-7(n2853)* mutant animals.

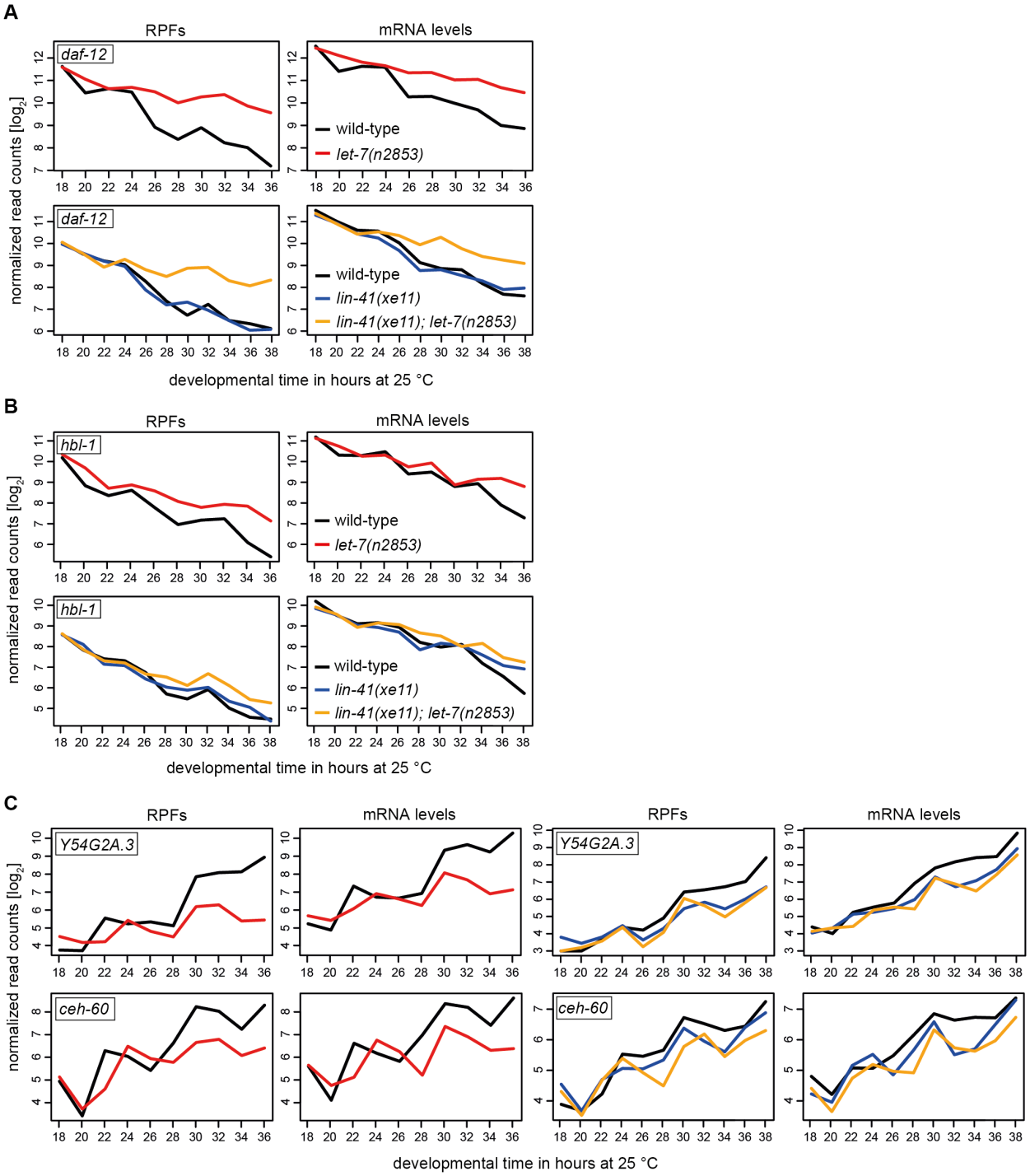


Figure S2. Related to Figures 1 and 2. Expression of *let-7* targets and candidate LIN41 targets over development
(A, B) Expression of the *let-7* targets (A) *daf-12* and (B) *hbl-1* over development. The columns on the left show ribosome-protected fragments (RPFs), the columns on the right show mRNA levels. Two independently performed time course experiments are depicted in the upper and lower panels, respectively.
(C) Expression of candidate LIN41 target genes over development. The four panels for each gene are arranged in rows, with RPF plots to the left of mRNA level plots for each time course experiment.

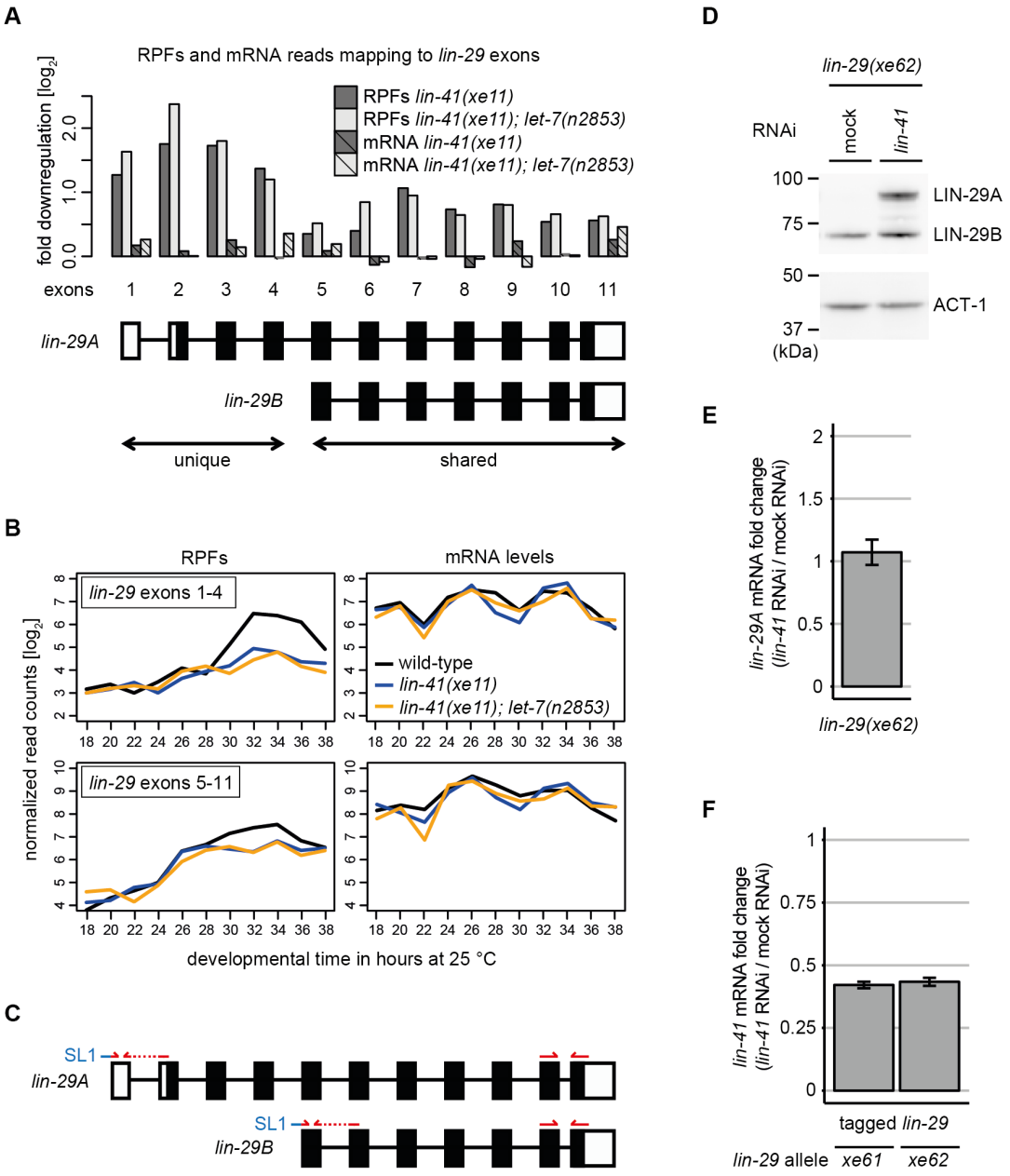
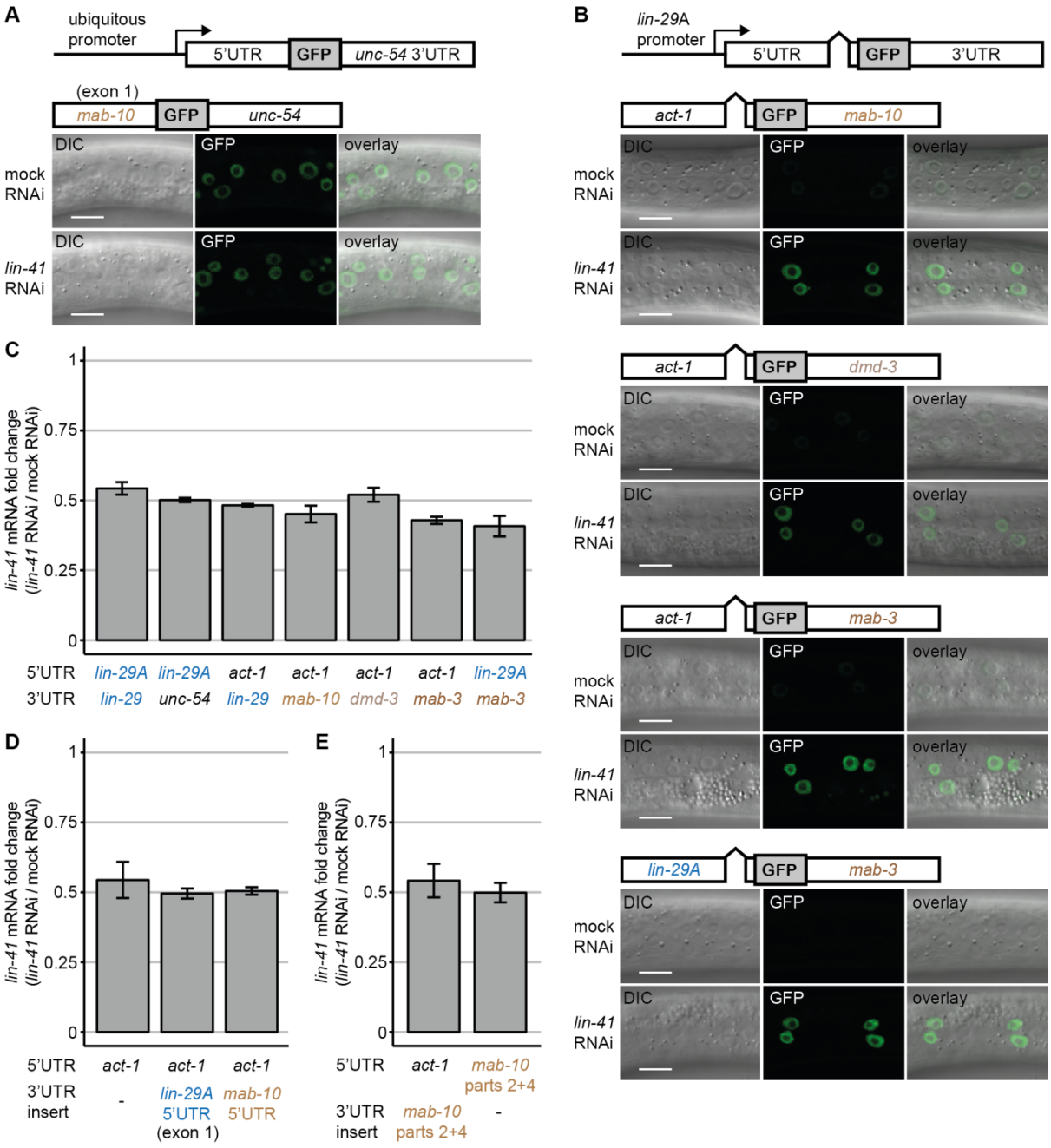


Figure S3. Related to Figure 4. LIN41 inhibits the translation of only one of two *lin-29* isoforms

(A) Quantification of the fold downregulation (log_2) in RPF and mRNA reads for each exon of the *lin-29* gene in the indicated mutants relative to wild-type worms. Reads were pooled from the five time points used for differential gene expression analysis in Figures 1D, E (28 - 36 h). Below the histogram is a schematic representation (not to scale) of the two *lin-29* isoforms according to a gene model from (Rougvie and Ambros, 1995); open boxes: UTRs, filled boxes: coding sequences, with the arrows indicating exons shared between the two isoforms or unique to *lin-29A*.

(B) Expression of *lin-29* over development in the indicated genetic backgrounds, separated by reads pooled from exons 1-4 and exons 5-11, respectively. Exons 1-4 are specific to *lin-29A*, exons 5-11 are shared between *lin-29A* and *lin-29B*.

(C) Schematic depiction of the RT-qPCR primers used in Figure 4C to individually measure the mRNA levels of *lin-29A* and *lin-29B*. Forward primers bind to the isoform-specific SL1 trans-splice leader junctions (Gu et al., 2012; Rougvie and Ambros, 1995), while reverse primers bind to the next downstream splice junction. The primers used to detect both isoforms at once anneal to exons 10 and 11, respectively.
(D) Western blot analysis to detect endogenous GFP::3xFLAG-tagged LIN-29A and LIN-29B proteins in *lin-29(xe62[lin-29::gfp::3xflag])* animals using an anti-FLAG antibody. The *lin-29(xe62[lin-29::gfp::3xflag])* allele was obtained independently from the *lin-29(xe61[lin-29::gfp::3xflag])* allele presented in Figure 4. ACT-1 was detected as a loading control. Animals were exposed to *lin-41* or mock RNAi and harvested as early L3 larval stage worms.
(E) RT-qPCR analysis to measure the fold change of *lin-29A* mRNA levels (normalized by *act-1* mRNA levels) in early L3 stage *lin-29(xe62[lin-29::gfp::3xflag])* animals exposed to *lin-41* RNAi compared to mock RNAi. n = 3 biological replicates, data as mean ± s.e.m.
(F) RT-qPCR analysis to measure the knock-down of *lin-41* mRNA in both *lin-29(xe61[lin-29::gfp::3xflag])* and *lin-29(xe62[lin-29::gfp::3xflag])* animals used for Western blot, RT-qPCR and confocal imaging analysis in (D, E) and Figure 4. Depicted is the fold change of *lin-41* mRNA levels (normalized by *act-1* mRNA levels) in worms grown on *lin-41* RNAi bacteria relative to those grown on mock RNAi bacteria. n = 3 biological replicates, data as mean ± s.e.m.



F

			repression on protein level	repression on mRNA level	conclusions
1.	*lin-29A*	*lin-29*	✓	✗	translational repression, no mRNA degradation
2.	*act-1*	*lin-29*	✗	✗	translational repression occurs through *lin-29A* 5'UTR
3.	*lin-29A*	*unc-54*	✓	✗	lack of mRNA degradation is not due to *lin-29* 3'UTR
4.	*act-1*	*mab-10*	✓	✓	there is degradative activity in the hypodermis
5.	*act-1*	*dmd-3*	✓	✓	
6.	*act-1*	*mab-3*	✓	✓	
7.	*lin-29A*	*mab-3*	✓	✓	lack of mRNA degradation is not due to *lin-29A* 5'UTR
8.	*act-1*	*unc-54*	✗	✗	LIN41 bound to the 3'UTR leads to mRNA degradation
9.	*act-1*	*unc-54* with *lin-29A* 5'UTR	✓	✓	
10	*act-1*	*unc-54* with *mab-10* 5'UTR	✗	✗	
11	*act-1*	*unc-54* with *mab-10* parts 2+4	✓	✓	LIN41 bound to the 5'UTR leads to translational repression
12	*mab-10* parts 2+4	*unc-54*	✓	✗	

Figure S4. Related to Figure 5. GFP reporter assays to study LIN41-mediated repression

(A, B) Micrographs of early L3 stage animals, exposed to *lin-41* or mock RNAi, expressing nuclear-localized GFP reporters with (A) the unregulated *unc-54* 3'UTR and the *mab-10* 5'UTR exon 1 sequence from a ubiquitously active *dpy-30* promoter or (B) the indicated 5'UTRs and 3'UTRs from the *lin-29A* promoter.

(C, D, E) The knock-down of *lin-41* mRNA in worm lines expressing nuclear-localized GFP reporters from the *lin-29A* promoter used in (B) and Figures 5A, 7B, 7D, and 7E was measured by RT-qPCR. Depicted is the fold change of *lin-41* mRNA levels (normalized by *act-1* mRNA levels) in early L3 stage worms grown on *lin-41* RNAi bacteria relative to those grown on mock RNAi bacteria. n = 3 biological replicates, data as mean ± s.e.m.

(F) Summary of the reporter experiments to study the different modes of LIN41 target silencing activity.

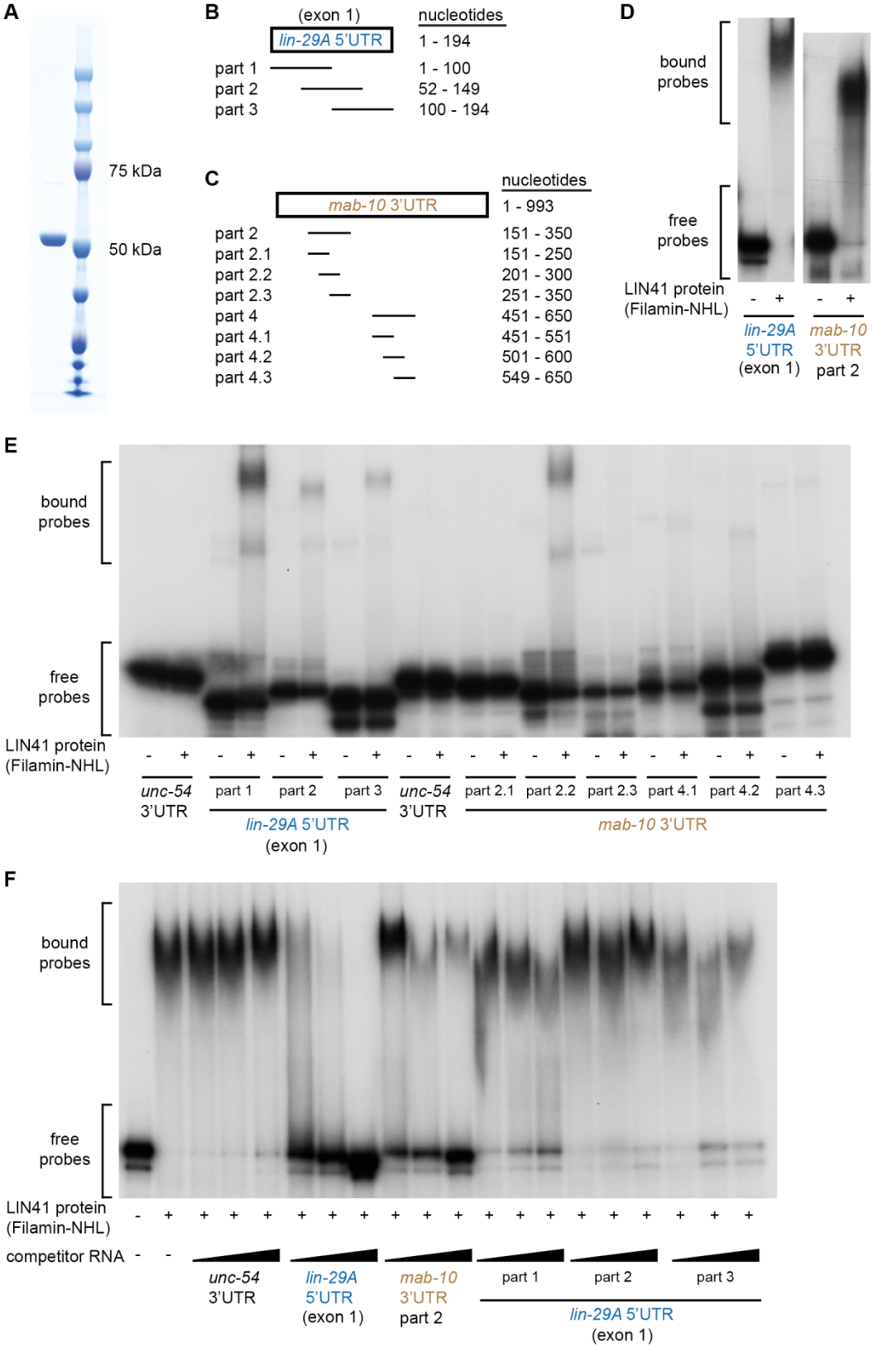

A
75 kDa
50 kDa
B
(exon 1)
nucleotides
lin-29A 5'UTR
1 - 194
part 1
1 - 100
part 2
52 - 149
part 3
100 - 194
C
nucleotides
mab-10 3'UTR
1 - 993
part 2
151 - 350
part 2.1
151 - 250
part 2.2
201 - 300
part 2.3
251 - 350
part 4
451 - 650
part 4.1
451 - 551
part 4.2
501 - 600
part 4.3
549 - 650
D
bound probes
free probes
LIN41 protein (Filamin-NHL)
- + - +
lin-29A 5'UTR (exon 1)
mab-10 3'UTR part 2
E
bound probes
free probes
LIN41 protein (Filamin-NHL)
unc-54 3'UTR
part 1
part 2
part 3
unc-54 3'UTR
part 2.1
part 2.2
part 2.3
part 4.1
part 4.2
part 4.3
lin-29A 5'UTR (exon 1)
mab-10 3'UTR
F
bound probes
free probes
LIN41 protein (Filamin-NHL)
competitor RNA
unc-54 3'UTR
lin-29A 5'UTR (exon 1)
mab-10 3'UTR part 2
part 1
part 2
part 3
lin-29A 5'UTR (exon 1)

Figure S5. Related to Figure 6. EMSAs to narrow down the region and test the specificity of LIN41 binding sites
(A) Coomassie-stained gel of the N-terminally truncated LIN41 protein fragment (containing the Filamin and NHL domains) used for EMSAs in Figures 6 and S5.
(B, C) Schematics of the radioactively labeled RNA gel-shift probes and cold competitor RNAs for EMSAs in (D, E, F) spanning parts of the *lin-29A* 5'UTR or *mab-10* 3'UTR.
(D) Autoradiographs showing binding of LIN41 to the 194 nt exon 1 fragment of *lin-29A* and the 200 nt- part 2 fragment from the *mab-10* 3'UTR.
(E) Autoradiograph to assay binding of LIN41 to partially overlapping ~100 nt fragments from the *lin-29A* 5'UTR or from parts of the *mab-10* 3'UTR. A 100 nt fragment from the *unc-54* 3'UTR is used as a negative control.
(F) Autoradiograph of competition EMSA with the *lin-29A* 5'UTR exon 1 (194 nt) as a radioactive probe. Different cold competitor RNAs were added to the reaction to test for competitive binding: whereas the *unc-54* 3'UTR fragment (198 nt) and the individual ~100 nt fragments of the *lin-29A* 5'UTR were unable to efficiently outcompete binding to the full-length *lin-29A* 5'UTR exon, the 200 nt part 2 of the *mab-10* 3'UTR and the full-length *lin-29A* 5'UTR exon (194 nt) interfered with binding to the radioactive probe moderately and strongly, respectively. Triangles represent 5-fold increases (from 12 nM to 300 nM) of cold competitor RNA.

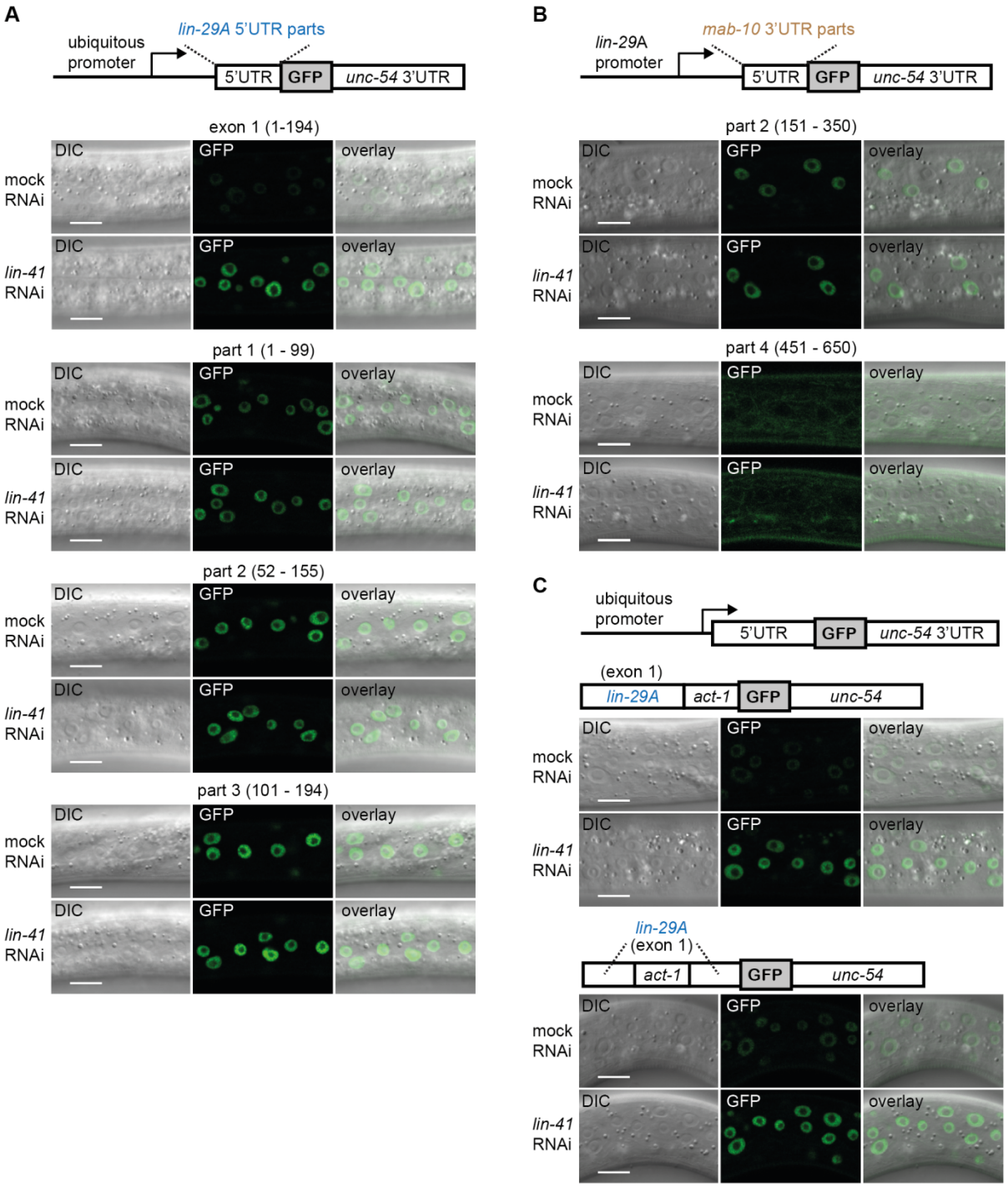


Figure S6. Related to Figures 6 and 7. GFP reporter assays to study LIN41-mediated repression at the 5'UTR
(A, B, C) Micrographs of early L3 stage animals, exposed to *lin-41* or mock RNAi, expressing nuclear-localized GFP reporters from (A, C) a ubiquitously active *dpy-30* promoter or (B) the *lin-29A* promoter. The reporters contain the unregulated *unc-54* 3'UTR and, as a 5'UTR, (A) the indicated *lin-29A* 5'UTR parts, (B) *mab-10* 3'UTR parts, or (C) the *lin-29A* 5'UTR exon 1, followed or centrally interrupted by an insertion of 65 nucleotides from the *act-1* 5'UTR.

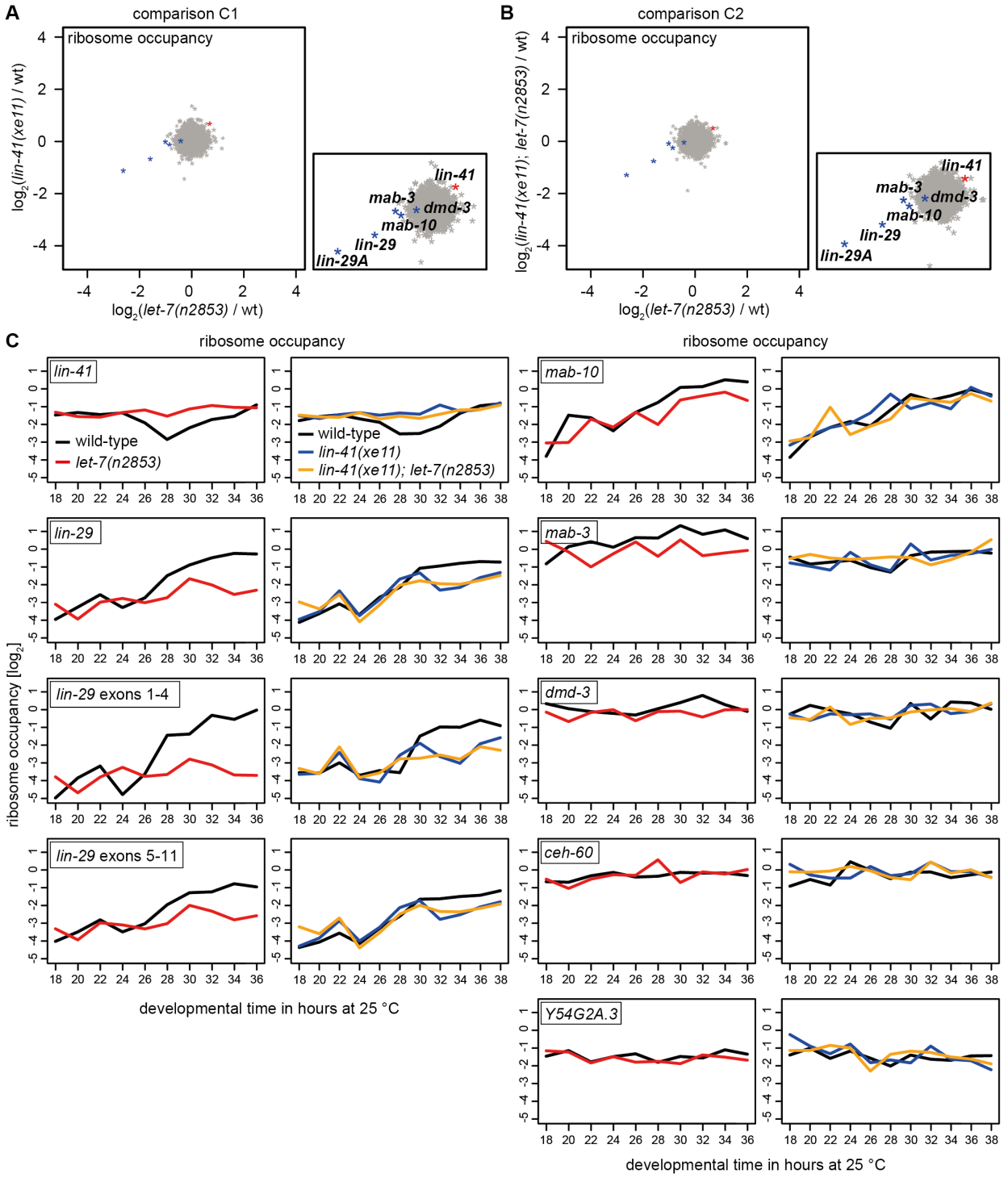

A
comparison C1
ribosome occupancy
log2(lin-41(xe11) / wt)
log2(let-7(n2853) / wt)
lin-41
mab-3
dmd-3
mab-10
lin-29
lin-29A
B
comparison C2
ribosome occupancy
log2(lin-41(xe11); let-7(n2853) / wt)
log2(let-7(n2853) / wt)
C
ribosome occupancy
wild-type
let-7(n2853)
lin-41(xe11)
lin-41(xe11); let-7(n2853)
lin-41
lin-29
lin-29 exons 1-4
lin-29 exons 5-11
mab-10
mab-3
dmd-3
ceh-60
Y54G2A.3
ribosome occupancy [log2]
developmental time in hours at 25 °C
D
RPF reads
wild-type
let-7(n2853)
32 h
34 h
36 h
lin-29A 5'
1 kb
E
RNA-seq reads

Figure S7. Related to Figure 7. Ribosome occupancies on LIN41 target mRNAs

(A, B) Scatter plots depicting mutant to wild-type fold changes (log_2) in ribosome occupancies (normalized RPF read counts divided by normalized RNA-seq read counts) for each gene. Ribosome occupancy changes in (A) *let-7(n2853)* and *lin-41(xe11)* or (B) *let-7(n2853)* and *lin-41(xe11); let-7(n2853)* mutant animals are compared. Asterisks representing *lin-41* and LIN41 targets are labeled in the zoom-in windows on the right of the scatter plots and colored red and blue, respectively. In addition, the ribosome occupancies for *lin-29A* were added to the plots, calculated by only considering RPF and RNA-seq counts mapping to *lin-29* exons 1-4. Each comparison (x-axis versus y-axis) is between two independent experiments, with independent wild-type replicates.

(C) Ribosome occupancies of *lin-41*, candidate and validated LIN41 target genes, as well as exons 1-4 and 5-11 of *lin-29* over development.

(D, E) Genome browser (IGV) screen shots visualizing the accumulation of RPF (D) or RNA-seq reads (E) mapping to the *lin-29A*-specific exons 1-4 in wild-type and *let-7(n2853)* mutant animals as indicated at different time points in development (32h, 34h, 36h). Reads are processed to wiggle files and depicted on a log_2 scale. Whereas RNA levels do not change (E), translational repression in *let-7(n2853)* mutant animals is observed as a decrease of RPF read coverage on each of the four *lin-29A*-specific exons, including the non-coding 5'UTR exon 1 (D). Open boxes: UTRs, filled boxes: coding sequences.

SUPPLEMENTAL TABLES

Table S1. Related to Figures 1 and 2. Expression data for the first time course experiment

WormBase ID, gene name and normalized $\log_2$ RPF and mRNA level read counts at each assayed time point of development are listed for all protein-coding genes, in wild-type N2 and *let-7(n2853)* mutant animals. Table S1 is provided as a Microsoft Excel file.

Table S2. Related to Figures 1 and 2. Expression data for the second time course experiment

WormBase ID, gene name and normalized $\log_2$ RPF and mRNA level read counts at each assayed time point of development are listed for all protein-coding genes, in wild-type N2, *lin-41(xe11)* mutant and *lin-41(xe11); let-7(n2853)* double mutant animals. Table S2 is provided as a Microsoft Excel file.

Table S3. Related to Figure 2. Genes differentially expressed in *let-7(n2853)* and *lin-41(xe11)* mutant animals

The gene names, WormBase IDs and earliest time point of dysregulation of genes differentially expressed in both *let-7(n2853)* and *lin-41(xe11)* mutant relative to wild-type animals are listed. Upregulated and downregulated genes are listed in two separate panels. Table S3 is provided as a Microsoft Excel file.

Table S4. Related to STAR Methods. *C. elegans* lines used in this study

Strain genotype	Source	Strain name
Wild-type	CGC	N2
let-7(n2853) X	(Reinhart et al., 2000) CGC	N/A
lin-41(xe11) I	Ecsedi et al., 2015	HW1329
lin-41(xe11) I; *let-7(n2853)* X	Ecsedi et al., 2015	HW1330
EG6699, *lin-41(n2914)* I; *xeSi197[Plin-41::flag::gfp::lin-41::lin-41 3'UTR, unc-119(+)]* II	This study	HW1574
EG6701, *xeSi55[Pdpy-30::sart-3::gfp::his::flag::xrn-2 3'UTR, unc-119(+)]* I	This study	HW1008
lin-29(xe61[lin-29::gfp::3xflag]) II	This study	HW1822
lin-29(xe62[lin-29::gfp::3xflag]) II	This study	HW1825
EG6699, *xeSi104[Pdpy-30::gfp(pest)-h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW874
EG6699, *xeSi193[Pdpy-30: gfp(pest)-h2b::mab-10 3'UTR, unc-119(+)]* II	This study	HW1568
EG6699, *xeSi246[Pdpy-30::gfp(pest)-h2b::mab-3 3'UTR, unc-119(+)]* II	This study	HW1796
EG6699, *xeSi247[Pdpy-30::gfp(pest)-h2b::dmd-3 3'UTR, unc-119(+)]* II	This study	HW1797
EG6699, *xeSi189[Pdpy-30::gfp(pest)/h2b::lin-29 3'UTR, unc-119(+)]* II	This study	HW1564
EG6699, *xeSi170[Pdpy-30::lin-29A 5'UTR (with intron)::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW1620
EG6699, *xeSi147[Pdpy-30::lin-29A 5'UTR exon 1::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW1619
EG6699, *xeSi259[Pdpy-30::mab-10 5'UTR::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW1831

EG6699, *xeSi251[Pdpy-30::mab-10 5'UTR exon 1::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW1812
EG6699, *xeSi249[Pdpy-30::lin-29A 5'UTR exon part 1::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW1810
EG6699, *xeSi250[Pdpy-30::lin-29A 5'UTR exon part 2::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW1811
EG6699, *xeSi248[Pdpy-30::lin-29A 5'UTR exon part 3::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW1809
EG6699, *xeSi344[Pdpy-30::lin-29A 5'UTR exon with act-1 5'UTR exon 3' insertion::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW2020
EG6699, *xeSi345[Pdpy-30:: lin-29A 5'UTR exon with act-1 5'UTR exon middle insertion::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW2021
EG6699, *xeSi194[Plin-29A::gfp(pest)/h2b::lin-29 3'UTR, unc-119(+)]* II	This study	HW1569
EG6699, *xeSi171[Plin-29A(act-1 5'UTR exon)::gfp(pest)/h2b::lin-29 3'UTR, unc-119(+)]* II	This study	HW1621
EG6699, *xeSi196[Plin-29A::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW1571
EG6699, *xeSi263[Plin-29A(act-1 5'UTR exon)::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW1837
EG6699, *xeSi265[Plin-29A(act-1 5'UTR exon)::gfp(pest)/h2b::mab-10 3'UTR, unc-119(+)]* II	This study	HW1839
EG6699, *xeSi288[Plin-29A(act-1 5'UTR exon)::gfp(pest)/h2b::dmd-3 3'UTR, unc-119(+)]* II	This study	HW1914
EG6699, *xeSi268[Plin-29A(act-1 5'UTR exon)::gfp(pest)/h2b::mab-3 3'UTR, unc-119(+)]* II	This study	HW1843
EG6699, *xeSi267[Plin-29A:: gfp(pest)/h2b::mab-3 3'UTR, unc-119(+)]* II	This study	HW1841
EG6699, *xeSi264[Plin-29A(act-1 5'UTR exon):: gfp(pest)/h2b::unc-54 3'UTR(lin-29A 5'UTR exon), unc-119(+)]* II	This study	HW1838
EG6699, *xeSi341[Plin-29A(act-1 5'UTR exon)::gfp(pest)/h2b::unc-54 3'UTR(mab-10 5'UTR), unc-119(+)]* II	This study	HW2017
EG6699, *xeSi303[Plin-29A(act-1 5'UTR exon)::gfp(pest)/h2b::unc-54 3'UTR(mab-10 3'UTR part 2), unc-119(+)]* II	This study	HW1954
EG6699, *xeSi305[Plin-29A(act-1 5'UTR exon)::gfp(pest)/h2b::unc-54 3'UTR(mab-10 3'UTR part 4), unc-119(+)]* II	This study	HW1956
EG6699, *xeSi306[Plin-29A(act-1 5'UTR exon)::gfp(pest)/h2b::unc-54 3'UTR(mab-10 3'UTR parts 2+4), unc-119(+)]* II	This study	HW1957
EG6699, *xeSi348[Plin-29Ashort::mab-10 3'UTR part 2::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW2024
EG6699, *xeSi349[Plin-29Ashort::mab-10 3'UTR part 4::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW2025
EG6699, *xeSi350[Plin-29Ashort::mab-10 3'UTR parts 2+4::gfp(pest)/h2b::unc-54 3'UTR, unc-119(+)]* II	This study	HW2026

Table S5. Related to STAR Methods. Plasmids used in this study

The plasmid names, backbone and insert, the cloning technique used to obtain the plasmid, and primer or gBlock sequences used for cloning are listed. Table S5 is provided as a Microsoft Excel file.

Table S6. Related to STAR Methods. qPCR primers used in this study

qPCR primer names, sequences and additional comments are listed.

qPCR primer name	Sequence	Comments
act-1 F1	gttgcccagaggctatgttc	
act-1 R1	caagagcggtgatttccttc	
GFP F1	cttgttgaattagatggtgatgtt	
GFP R1	acaagtgttggccatgga	
lin-41 F2	acatcctggaaagcatcgag	
lin-41 R2	aagcgttgacgtgtgtatcg	
unc-54 F1	ctgctatgctcatctacacct	
unc-54 R1	tgtggtggcatttctgtctt	
mab-10 F1	tctccgatttttgagtcagctgt	
mab-10 R1	gagaacttgaacgccaacgg	

mab-3 F1	acagaaatcccgagatggtaaaga	
mab-3 R1	ggacttgctgatgttccaattatct	
dmd-3 F1	ccgtcgccgatagatacagt	
dmd-3 R1	gttgggcacacttcagacac	
Y54G2A.3 F1	atggatccgtacaaactagaagc	
Y54G2A.3 R1	tccgttgcattttctttgctct	
ceh-60 F1	ttgacaaaaacacgacggac	
ceh-60 R1	ttccaaaccaattgttaacctgt	
lin-29 F1	ccgacgagtacgaagaatgg	detects both isoforms (primers anneal to exons 10 and 11)
lin-29 R1	gtgattgtgggttgaacacg	
lin-29 F4[a]	ccagcacatcattcgatcact	*lin-29A* specific detection (exon3-exon4 junction)
lin-29 R4[a]	gaagttcagtagatccgcttga	
lin-29 F5[b]	acccaagtttgagttcgaaca	*lin-29B* specific detection (SL1-exon5 junction)
lin-29 R5[b]	gatgagttggcaaatgccttga	
lin-29 F6[b]	tacccaagtttgaggttccga	*lin-29A* specific detection (SL1-exon1 junction)
lin-29 R6[b]	tcgcacagtttcaggttacc	

[a]used to measure fold change of tagged *lin-29A*.
[b]used to measure fold enrichments in LIN-41 RNA-IPs.

Table S7. Related to STAR Methods. Primers used to obtain PCR templates for EMSA probes

PCR products used as templates for in-vitro transcription, primer names and primer sequences are listed.

PCR product	PCR primer name	Sequence[c]
unc-54 3'UTR fragment (198 nt)	unc-54_3UTR_F	AATTAACCCTCACTAAAGGGAGAAtaagtccaattactcttcaacatcc
	unc-54_3UTR_R	ggagggagcacaatttttttcg
lin-29A 5'UTR exon (194 nt)	lin-29A_5UTR_F	AATTAACCCTCACTAAAGGGAGAAgttccgaaccaattcgcctag
	lin-29A_5UTR_R	aatatttctgctccaaattcgcaca
mab-10 5'UTR (244 nt)	mab-10_5UTR_F	AATTAACCCTCACTAAAGGGAGAAaccccggcctcactgttc
	mab-10_5UTR_R	gatgccttctgaaggcacaca
mab-10 3'UTR part 1 (200 nt)	mab-10_3UTR_F1	AATTAACCCTCACTAAAGGGAGAAagctcggatttccatatttttg
	mab-10_3UTR_R1	taggaagagaaggaaattatatgactcc
mab-10 3'UTR part 2 (200 nt)	mab-10_3UTR_F2	AATTAACCCTCACTAAAGGGAGAAtaaattcccgccccctc
	mab-10_3UTR_R2	tatttaaattatctaattgggggcgg
mab-10 3'UTR part 3 (200 nt)	mab-10_3UTR_F3	AATTAACCCTCACTAAAGGGAGAAgttccaaaaccaacgccca
	mab-10_3UTR_R3	gagacccaaaaatttataaaagcaatc
mab-10 3'UTR part 4 (200 nt)	mab-10_3UTR_F4	AATTAACCCTCACTAAAGGGAGAAtcgcatccaatgcattgca
	mab-10_3UTR_R4	ggattattcaaatttttggagcaa
mab-10 3'UTR part 5 (200 nt)	mab-10_3UTR_F5	AATTAACCCTCACTAAAGGGAGAAtttccggtcattcgaaatgg
	mab-10_3UTR_R5	ccatgtaaacatttttggggaga
mab-10 3'UTR part 6 (200 nt)	mab-10_3UTR_F6	AATTAACCCTCACTAAAGGGAGAAactagtctcctccacccat
	mab-10_3UTR_R6	cccatctacttttttgtttccag
unc-54 3'UTR fragment (100 nt)	unc-54_3UTR_F	AATTAACCCTCACTAAAGGGAGAAcccaccccctatttttgttattat
	unc-54_3UTR_R	ttgaatctacacaatttcattgttagag
lin-29A 5'UTR part 1 (100 nt)	lin-29A_5UTR_F	AATTAACCCTCACTAAAGGGAGAAgttccgaaccaattcgcctag
	lin-29A_5UTR_R1	aaatattggaactataatcaaactattc
lin-29A 5'UTR part 2 (98 nt)	lin-29A_5UTR_F2	AATTAACCCTCACTAAAGGGAGAActtttctgaaaagagcctact
	lin-29A_5UTR_R2	aagccttggttccagtattag
lin-29A 5'UTR part 3 (95 nt)	lin-29A_5UTR_F3	AATTAACCCTCACTAAAGGGAGAAcgatcaaatcgacacacctacta
	lin-29A_5UTR_R	aatatttctgctccaaattcgcaca
mab-10 3'UTR part 2.1 (100 nt)	mab-10_3UTR_F2	AATTAACCCTCACTAAAGGGAGAAtaaattcccgccccctc
	mab-10_3UTR_R2b	aaaaccgccggaaaaatg
mab-10 3'UTR part 2.2 (100 nt)	mab-10_3UTR_F2c	AATTAACCCTCACTAAAGGGAGAAtcacttttttgtctagttttca
	mab-10_3UTR_R2c	cgaactgtgtttagactc
mab-10 3'UTR part 2.3 (100 nt)	mab-10_3UTR_F2b	AATTAACCCTCACTAAAGGGAGAAtcaaaaaccaatgccaagc
	mab-10_3UTR_R2	tatttaaattatctaattgggggcgg
mab-10 3'UTR part 4.1 (101 nt)	mab-10_3UTR_F4	AATTAACCCTCACTAAAGGGAGAAtcgcatccaatgcattgca
	mab-10_3UTR_R4b	cactgcattaaatgcaaa
mab-10 3'UTR part 4.2 (100 nt)	mab-10_3UTR_F4c	AATTAACCCTCACTAAAGGGAGAAgaagcgaaaagcccatag
	mab-10_3UTR_R4c	atttgaaaatttgatctattcag
mab-10 3'UTR part 4.3 (102 nt)	mab-10_3UTR_F4b	AATTAACCCTCACTAAAGGGAGAAgtgaaaattcaaaaaattcaaag
	mab-10_3UTR_R4	ggattattcaaatttttggagcaa

Table S3

Gene name	WormBase ID	hours of development when first upregulated consistently in *let-7(n2853)*, *lin-41(xe11)* and *lin-41(xe11); let-7(n2853)* mutant worms, compared to wild-type worms (on RPF level)	comment
lin-41	WBGene00003026	24	
col-90	WBGene00000665	32	
F41E6.14	WBGene00018295	32	
F52E1.5	WBGene00018694	32	
T20D4.10	WBGene00020616	32	
Y48E1B.8	WBGene00013007	32	
nhr-25	WBGene00003623	34	
col-17	WBGene00000606	34	
col-54	WBGene00000631	34	
F54A5.1	WBGene00018786	34	
col-156	WBGene00000729	34	
dpy-1	WBGene00001063	34	
R08E3.4	WBGene00019960	34	
T06D8.1	WBGene00011522	34	
col-41	WBGene00000618	34	
daf-9	WBGene00000905	34	
T19B10.5	WBGene00011833	34	
T20D4.12	WBGene00020618	34	
F46F2.3	WBGene00009787	34	
ifa-3	WBGene00002051	34	
F18E9.3	WBGene00017569	36	
K01D12.9	WBGene00010468	36	
ptr-13	WBGene00004227	36	
T06E4.10	WBGene00011537	36	
ins-4	WBGene00002087	36	lin-41 independent change from 32 hours on
pgp-5	WBGene00003999	-	inconsistant change, oscillating gene
C50D2.1	WBGene00016805	-	inconsistant change, oscillating gene
D2096.1	WBGene00017069	-	inconsistant change
col-98	WBGene00000673	-	inconsistant change, oscillating gene
grl-3	WBGene00001712	-	inconsistant change
col-142	WBGene00000715	-	inconsistant change, oscillating gene
pgp-6	WBGene00004000	-	inconsistant change, oscillating gene

Gene name	WormBase ID	hours of development when first downregulated consistently in *let-7(n2853)*, *lin-41(xe11)* and *lin-41(xe11); let-7(n2853)* mutant worms compared to wild-type worms (on RPF level)	comment
ceh-60	WBGene00017690	28	
mab-10	WBGene00003107	28	
lin-29	WBGene00003015	30	
dmd-3	WBGene00012832	30	
mab-3	WBGene00003100	30	
Y54G2A.3	WBGene00021869	30	
col-175	WBGene00000748	32	
col-138	WBGene00000711	32	
C53B4.8	WBGene00008277	32	
E01G4.6	WBGene00008448	32	
col-38	WBGene00000615	32	
F58E6.13	WBGene00077697	32	
bli-1	WBGene00000251	32	
bli-2	WBGene00000252	32	
col-49	WBGene00000626	32	
col-71	WBGene00000647	32	
rol-1	WBGene00004394	32	

col-112	WBGene00000686	32	
dao-4	WBGene00000930	32	
ZK105.1	WBGene00022653	32	
col-63	WBGene00000639	32	
F48B9.5	WBGene00018591	34	
col-104	WBGene00000678	34	
W03D2.9	WBGene00020983	34	
col-109	WBGene00000683	34	
col-48	WBGene00000625	34	
col-97	WBGene00000672	34	
B0034.5	WBGene00015005	34	
C04G6.2	WBGene00015453	34	
col-148	WBGene00000721	34	
col-124	WBGene00000698	34	
C29A12.6	WBGene00007800	34	
col-178	WBGene00000751	34	
C42D4.13	WBGene00016598	34	
col-14	WBGene00000603	34	
col-119	WBGene00000693	34	
col-20	WBGene00000609	34	
col-184	WBGene00000757	34	
col-60	WBGene00000636	34	
col-140	WBGene00000713	34	
col-152	WBGene00000725	34	
col-81	WBGene00000657	34	
grl-27	WBGene00001736	34	
col-139	WBGene00000712	34	
vit-3	WBGene00006927	34	
vit-6	WBGene00006930	34	lin-41 independent change from 32 hours on
col-129	WBGene00000703	34	
R09E10.6	WBGene00011176	34	
R17.3	WBGene00011269	34	
col-122	WBGene00000696	34	
T10E9.3	WBGene00020413	34	
T21B10.6	WBGene00011888	34	
col-181	WBGene00000754	34	
col-88	WBGene00000663	34	
col-120	WBGene00000694	34	
col-106	WBGene00000680	34	
C44B12.1	WBGene00016636	34	
T26H5.9	WBGene00044206	34	
clec-197	WBGene00008202	34	
K01D12.1	WBGene00010461	34	
vit-2	WBGene00006926	34	lin-41 independent change from 30 hours on
C45B2.1	WBGene00016658	36	
vit-4	WBGene00006928	36	
clec-229	WBGene00010355	36	
T06C12.8	WBGene00011515	36	
col-137	WBGene00000710	36	
col-19	WBGene00000608	36	
col-179	WBGene00000752	36	
Y41C4A.11	WBGene00012757	36	
col-8	WBGene00000597	36	

Table S5

Short name	Plasmid	Insert	Cloning technique	Primers[a] (fwd, rev) or gBlocks[b]	
pHB1	pENTR L4-R1	*lin-29A* promoter	BP reaction	fwd	ggggacaactttgtatagaaaagttgggGATATATTTTGATCGCTACTCAACA
				rev	ggggactgctttttttgtacaaacttggTGCGTTGAAGAAGTTGGCTTGA
pFA1	pENTR L4-R1	*lin-29A* promoter with *act-1* 5'UTR exon	Gibson assembly[c]	fwd(1)	CTCCTTACGGGTTTTTTGTC
				rev(1)	cttttcgaagctattttcttctcaagaatatatccttggaaagccCTGCAACAATTGTAACTTTTAG
				fwd(2)	gagaagaaaatagcttcgaaaaggatagacaaaaccggtcaatGTGAGTTTGTTTGCCTGCCT
				rev(2)	ACCATGTAATACGACTCACT
pFA181	pENTR L4-R1	*lin-29A* short promoter	Gibson assembly[d]	gBlock	ctccttacgggtttttgtctcgatATCAACGTTGCTATTTTTTTGCGATAGGTTTTTTTCCT TTCTAAAAGTTACAATTGTTGCAGCAAGTTTGTACAAAAAAGTTGAACGAGAA ACGTAAAATGATATAAATATCAATATATTAAATTAGATTTTGCATAAAAAACAG ACTACATAATACTGTAAAACACAACATATGCAGTCACTATGAATCAACTACTT AGATGGTATTAGTGACCTGTAGAATTCGAGCTCTAGAGCTGCAGGGCGGCC GCGATatcccctatagtgagtcgtattacatggt
(1)	pENTR L4-R1	*dpy-30* promoter	pENTR_L4-R1_Pdpy-30, published in (Ecsedi et al., 2015)		
pHB2	pENTR L4-R1	*lin-41* promoter	BP reaction	fwd	ggggacaactttgtatagaaaagttgTACCACGCAGACAAGGAGCTAC
				rev	ggggactgctttttgtacaaacttggCACTTTTTCCAAGTCTGAAAAGG
(2)	pENTR L1-L2	GFP(PEST)-H2B	pBMF2.7, published in (Wright et al., 2011)		
pFA2	pENTR L1-L2	*lin-29* 5'UTR (whole, with intron) + GFP(PEST)-H2B	Gibson assembly[e]	fwd(1)	ccaagtttgtacaaaaaagcaggctccGTTCCGAACCAATTCGCCTAG
				rev(1)	agccatggctaagtctagacatTGCGTTGAAGAAGTTGGCT
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pFA3	pENTR L1-L2	*lin-29* 5'UTR exon 1 + GFP(PEST)-H2B	Gibson assembly[e]	fwd(1)	ccaagtttgtacaaaaaagcaggctccGTTCCGAACCAATTCGCCTAG
				rev(1)	agccatggctaagtctagacatAATATTTCTGCTCCAAATTCGCAC
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pFA33	pENTR L1-L2	*lin-29* 5'UTR exon 1 part 1 + GFP(PEST)-H2B	Gibson assembly[f]	gBlock	tgccaagtttgtacaaaaaagcaggctccGTTCCGAACCAATTCGCCTAGCGAACCTTGG AGTCCAACTCCGTTTTTACTCTTTTCTGAAAAGAGCCTACTAAATATTGGAAC TATAATCAAACTATTatgtctagacttagccatggcttc
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pFA34	pENTR L1-L2	*lin-29* 5'UTR exon 1 part 2 + GFP(PEST)-H2B	Gibson assembly[f]	gBlock	tgccaagtttgtacaaaaaagcaggctccCTTTTCTGAAAAGAGCCTACTAAATATTGGAA CTATAATCAAACTATTCGATCAAATCGACACACCTACTAGATATCAAATATCT AATACTGGAACCAAGGCTTatgtctagacttagccatggcttc
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pFA35	pENTR L1-L2	*lin-29* 5'UTR exon 1 part 3 + GFP(PEST)-H2B	Gibson assembly[f]	gBlock	tgccaagtttgtacaaaaaagcaggctccGATCAAATCGACACACCTACTAGATATCAAAT ATCTAATACTGGAACCAAGGCTTGGGTAACCTGAAACTGTGCGAATTTGGA GCAGAAATATTatgtctagacttagccatggcttc
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pFA177	pENTR L1-L2	*lin-29* 5'UTR exon 1 with 3' act-1 5'UTR insertion + GFP(PEST)-H2B	Gibson assembly[f]	gBlock	tgccaactttgtacaaaaaagcaggctccGTTCCGAACCAATTCGCCTAGCGAACCTTGG AGTCCAACTCCGTTTTTACTCTTTTCTGAAAAGAGCCTACTAAATATTGGAAC TATAATCAAACTATTCGATCAAATCGACACACCTACTAGATATCAAATATCTA ATACTGGAACCAAGGCTTGGGTAACCTGAAACTGTGCGAATTTGGAGCAGA AATATTGGCTTTCCAAGGATATATTCTTGAGAAGAAAATAGCTTCGAAAAGG ATAGACAAAACCGGTCAATatgtctagacttagccatggcttc
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pFA183	pENTR L1-L2	*lin-29* 5'UTR exon 1 with middle act-1 5'UTR insertion + GFP(PEST)-H2B	Gibson assembly[f]	gBlock	tgccaactttgtacaaaaaagcaggctccGTTCCGAACCAATTCGCCTAGCGAACCTTGG AGTCCAACTCCGTTTTTACTCTTTTCTGAAAAGAGCCTACTAAATATTGGAAC TATAATCAAACTATTGGCTTTCCAAGGATATATTCTTGAGAAGAAAATAGCTT CGAAAAGGATAGACAAAACCGGTCAATCGATCAAATCGACACACCTACTAG ATATCAAATATCTAATACTGGAACCAAGGCTTGGGTAACCTGAAACTGTGCG AATTTGGAGCAGAAATATTatgtctagacttagccatggcttc
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pFA4	pENTR L1-L2	*mab-10* 5'UTR (whole, without intron) + GFP(PEST)-H2B	Gibson assembly[f]	gBlock	tgccaactttgtacaaaaaagcaggctccACCCCGGCCTCACTGTTCCACAGCCTCCTGA CCCGGTCGTGCCCGGACCTTGTGCGGCGCCAGACGCGAATGATATTCGGA GACACGGGCTGAATGGTTTCGTCGGTAGAGAACCACGTCATCCCGTTTTCG TCGTCTGCGGAGTTAGAACTGTATCAGGATACTGAGGCTATTCGATAGCCG CTGCCCTTCCGTCGATCTGGATTGGGAGCGTACCTCAACTTGTGTGCCTTC AGAAGGCATCatgtctagacttagccatggcttc
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pFA5	pENTR L1-L2	*mab-10* 5'UTR exon 1 + GFP(PEST)-H2B	Gibson assembly[f]	gBlock	tgccaactttgtacaaaaaagcaggctccACCCCGGCCTCACTGTTCCACAGCCTCCTGA CCCGGTCGTGCCCGGACCTTGTGCGGCGCCAGACGCGAATGATATTCGGA GACACGGGCTGAATGGTTTCGTCGGTAGAGAACCACGTCATCCCGTTTTCG TCGTCTGCGGAGTTAGAACTGTATCAGatgtctagacttagccatggcttc
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pFA81	pENTR L1-L2	*mab-10* 3'UTR part 2 +	Gibson assembly[f]	gBlock	tgccaactttgtacaaaaaagcaggctccTAAATTCCCGCCCCCTTCGCTCGGAGTCATA TAATTTCCTTCTCTTCCTATCACTTTTTTGTCTAGTTTTCAAAAAACCCCCCAT TTTTCCGGCGGTTTTTCAAAAACCAATGCCAAGCTCTCATATAATGGGAGTC

		GFP(PEST)-H2B			TAAACACAGTTCGGTTCCAAAACCAACGCCCACGGCCCCGCCCCCAATTAG ATAATTTAAATAatgtctagacttagccatggcttc
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pFA178	pENTR L1-L2	*mab-10* 3'UTR part 4 + GFP(PEST)-H2B	Gibson assembly[f]	gBlock	tgccaactttgtacaaaaaagcaggctccTCGCATCCAATGCATTGCAAGTTGATTGCTTT TATAAATTTTTGGGTCTCGAAGCGAAAAGCCCATAGGATTTTCACAAAAAATTT TGCATTTAATGCAGTGAAAATTCAAAAAATTCAAAGATTTTTCTGAATAGATC AAATTTTCAAATTTTCCGGTCATTCGAAATGGCCTAATTTGCTCCAAAAATTT GAATAATCCatgtctagacttagccatggcttc
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pFA179	pENTR L1-L2	*mab-10* 3'UTR parts 2+4 + GFP(PEST)-H2B	Gibson assembly[f]	gBlock	tgccaactttgtacaaaaaagcaggctccTAAATTCCCGCCCCCTTCGCTCGGAGTCATA TAATTTCCTTCTCTTCCTATCACTTTTTTGTCTAGTTTTCAAAAAACCCCCCAT TTTTCCGGCGGTTTTTCAAAAACCAATGCCAAGCTCTCATATAATGGGAGTC TAAACACAGTTCGGTTCCAAAACCAACGCCCACGGCCCCGCCCCCAATTAG ATAATTTAAATATCGCATCCAATGCATTGCAAGTTGATTGCTTTTATAAATTTT TGGGTCTCGAAGCGAAAAGCCCATAGGATTTTCACAAAAATTTTGCATTTAA TGCAGTGAAAATTCAAAAAATTCAAAGATTTTTCTGAATAGATCAAATTTTCA AATTTTCCGGTCATTCGAAATGGCCTAATTTGCTCCAAAAATTTGAATAATCC atgtctagacttagccatggcttc
				fwd(2)	agccaacttcttcaacgcaATGTCTAGACTTAGCCATGGCT
				rev(2)	TGGGACAACTCCAGTGAAAAG
pHB3	pENTR L1-L2	FLAG::GFP::LIN41	Fusion PCR + BP reaction[g]	fwd(1)	**ggtagcggcagcggtagc**GCGACCATCGTGCCATG
				rev(1)	ggggaccactttgtacaagaaagctgggtcCTAGAAGACACGGATGCAATT
				fwd(2)	ggggacaagtttgtacaaaaaagcaggctccATGGACTACAAAGACGATGACGA
				rev(2)	**gctaccgctgccgctacc**AGCTGGGTCTGAAAATACAGGTT
(3)	pENTR R2-L3	*unc-54* 3'UTR	pCM5.37, published in (Merritt et al., 2008)		
pHB4	pENTR R2-L3	*lin-29* 3'UTR	BP reaction	fwd	ggggacagctttcttgtacaaagtggTAATTTTAATTTTTTTTTGAATTTTTTCTAA
				rev	ggggacaactttgtataataaagttgATACATAATCGTTTATATTTTCAATC
pFA6	pENTR R2-L3	*mab-10* 3'UTR	BP reaction	fwd	ggggacagctttcttgtacaaagtggATCTTGAAGCTCGGATTTCCAT
				rev	ggggacaactttgtataataaagttgTGTTACGGGAATCATGTCTTC
pFA7	pENTR R2-L3	*mab-3* 3'UTR	BP reaction	fwd	ggggacagctttcttgtacaaagtggTAAGATCTATAATTTTGACCAATTAT
				rev	ggggacaactttgtataataaagttgCGTGGAGCAGAACGTCTC
pFA8	pENTR R2-L3	*dmd-3* 3'UTR	BP reaction	fwd	ggggacagctttcttgtacaaagtggAAACTCTAAAATAGTTTGAATTTTTAAATT
				rev	ggggacaactttgtataataaagttgCCCGAAGTGTCAGCCTATATT
pFA9	pENTR R2-L3	*unc-54* 3'UTR with *lin-29* 5'UTR exon 1	Gibson assembly[h]	gBlock	acttctcttaatttctttgtggGTTCCGAACCAATTCGCCTAGCGAACCTTGGAGTCCAAC TCCGTTTTTACTCTTTTCTGAAAAGAGCCTACTAAATATTGGAACTATAATCA AACTATTCGATCAAATCGACACACCTACTAGATATCAAATATCTAATACTGGA ACCAAGGCTTGGGTAACCTGAAACTGTGCGAATTTGGAGCAGAAATATTccttt ttagcttcttttaagtcac
pFA64	pENTR R2-L3	*unc-54* 3'UTR with *mab-10* 5'UTR	Gibson assembly[h]	gBlock	acttctcttaatttctttgtggACCCCGGCCTCACTGTTCCACAGCCTCCTGACCCGGTC GTGCCCGGACCTTGTGCGGCGCCAGACGCGAATGATATTCGGAGACACGG GCTGAATGGTTTCGTCGGTAGAGAACCACGTCATCCCGTTTTCGTCGTCTG CGGAGTTAGAACTGTATCAGGATACTGAGGCTATTCGATAGCCGCTGCCCT TCCGTCGATCTGGATTGGGAGCGTACCTCAACTTGTGTGCCTTCAGAAGGC ATCccttttagcttctttaagtcac
pFA90	pENTR R2-L3	*unc-54* 3'UTR with *mab-10* 3'UTR part 2	Gibson assembly[h]	gBlock	acttctcttaatttctttgtggTAAATTCCCGCCCCCTTCGCTCGGAGTCATATAATTTCCT TCTCTTCCTATCACTTTTTTGTCTAGTTTTCAAAAAACCCCCCATTTTTCCGG CGGTTTTTCAAAAACCAATGCCAAGCTCTCATATAATGGGAGTCTAAACACA GTTCGGTTCCAAAACCAACGCCCACGGCCCCGCCCCCAATTAGATAATTTA AATAccttttagcttctttaagtcac
pFA92	pENTR R2-L3	*unc-54* 3'UTR with *mab-10* 3'UTR part 4	Gibson assembly[h]	gBlock	acttctcttaatttctttgtggTCGCATCCAATGCATTGCAAGTTGATTGCTTTTATAAATT TTTGGGTCTCGAAGCGAAAAGCCCATAGGATTTTCACAAAAAATTTTGCATTT AATGCAGTGAAAATTCAAAAAATTCAAAGATTTTTCTGAATAGATCAAATTTT CAAATTTTCCGGTCATTCGAAATGGCCTAATTTGCTCCAAAAATTTGAATAAT CCccttttagcttctttaagtcac
pFA93	pENTR R2-L3	*unc-54* 3'UTR with *mab-10* 3'UTR parts 2+4	Gibson assembly[h]	gBlock	acttctcttaatttctttgtggTAAATTCCCGCCCCCTTCGCTCGGAGTCATATAATTTCCT TCTCTTCCTATCACTTTTTTGTCTAGTTTTCAAAAAACCCCCCATTTTTCCGG CGGTTTTTCAAAAACCAATGCCAAGCTCTCATATAATGGGAGTCTAAACACA GTTCGGTTCCAAAACCAACGCCCACGGCCCCGCCCCCAATTAGATAATTTA AATATCGCATCCAATGCATTGCAAGTTGATTGCTTTTATAAATTTTTGGGTCT CGAAGCGAAAAGCCCATAGGATTTTCACAAAAAATTTTGCATTTAATGCAGTG AAAATTCAAAAAATTCAAAGATTTTTCTGAATAGATCAAATTTTCAAATTTTCC GGTCATTCGAAATGGCCTAATTTGCTCCAAAAATTTGAATAATCCccttttagcttct tttaagtcac
(4)	pENTR R2-L3	*lin-41* 3'UTR	pENTR_R2-L3_lin-41 3'UTR, published in (Ecsedi et al., 2015)		
pFA10	pCFJ150	(1), (2), (3)	LR reaction		
pFA11	pCFJ150	(1), (2), pFA6	LR reaction		
pFA12	pCFJ150	(1), (2), pFA7	LR reaction		
pFA13	pCFJ150	(1), (2), pFA8	LR reaction		
pHB7	pCFJ150	(1), (2), pHB4	LR reaction		
pFA15	pCFJ150	(1), pFA2, (3)	LR reaction		
pFA16	pCFJ150	(1), pFA3, (3)	LR reaction		
pFA17	pCFJ150	(1), pFA4, (3)	LR reaction		
pFA18	pCFJ150	(1), pFA5, (3)	LR reaction		
pFA52	pCFJ150	(1), pFA33, (3)	LR reaction		
pFA53	pCFJ150	(1), pFA34, (3)	LR reaction		
pFA54	pCFJ150	(1), pFA35, (3)	LR reaction		
pFA186	pCFJ150	(1), pFA177, (3)	LR reaction		

pFA187	pCFJ150	(1), pFA183, (3)	LR reaction		
pHB5	pCFJ150	pHB1, (2), pHB4	LR reaction		
pHB8	pCFJ150	pHB1, (2), (3)	LR reaction		
pFA20	pCFJ150	pFA1, (2), pHB4	LR reaction		
pFA21	pCFJ150	pFA1, (2), pFA6	LR reaction		
pFA22	pCFJ150	pFA1, (2), pFA7	LR reaction		
pFA67	pCFJ150	pFA1, (2), pFA8	LR reaction		
pFA23	pCFJ150	pHB1, (2), pFA7	LR reaction		
pFA24	pCFJ150	pFA1, (2), (3)	LR reaction		
pFA25	pCFJ150	pFA1, (2), pFA9	LR reaction		
pFA175	pCFJ150	pFA1, (2), pFA64	LR reaction		
pFA96	pCFJ150	pFA1, (2), pFA90	LR reaction		
pFA98	pCFJ150	pFA1, (2), pFA92	LR reaction		
pFA99	pCFJ150	pFA1, (2), pFA93	LR reaction		
pFA190	pCFJ150	pFA181, pFA81, (3)	LR reaction		
pFA191	pCFJ150	pFA181, pFA178, (3)	LR reaction		
pFA192	pCFJ150	pFA181, pFA179, (3)	LR reaction		
pHB6	pCFJ150	pHB2, pHB3, (4)	LR reaction		
pFA26	pDD282	regions for homologous recombination at the *lin-29* C-terminus, including a flexible linker sequence (underlined) and three silent mutations (bold underlined)	Gibson assembly[i]	gBlock1	agtcacgacgttgtaaaacgacggccagtcgccggcaCGAAATGATGACAACGAGGGAAATT TCGTAAATCGACACAAATCGACTCTTCCGCGCGCACCGCCCCATTGGGTCC CGGATGTCTGTCACTCTCTAACTATCCACACTCTCTAATATAAACGTTTAAAT TTACGATTTTGCCACCAATACTAACCAGCATTAGAGCGCGATTGCGCCCCAA TTTCAAAAGGTGCATTTTTTGCGATTGTATAATAAAACATTGTATTTTTTTTTC AAATCCAGGTTCAACATGATAACCCCACTGGAGAACATCCAACGCTACAAC GGGTCATCTTCCTCGGCCACGGCTGTCGTAACGGCGACCGGCTCGGCCGT CGTCTCATCGACACCCTCATCCACGTCATCCTCGTCAGCAGGTTCGTCCTC AAGTCAGGGAGGCGTGTTCAACCCACA**G**TC**T**CT**C**ATAAATAATATGAAAAAT CATTCCTATGGAGCATCGGGAGCCTCAGGAGCATCGatgagtaaaggagaagaattg ttcactggagttgtccca
				gBlock2	cgacgacaagcgtgattacaaggatgacgatgacaagagaTAATTTTAATTTTTTTTTGAATTT TTTCTAATTCCTGAACTCACATATTCCGATAAAATATTCATTTTTGGGGCCAA GCCCATTTCTCCAATTGGAATAAGATCTTCGTCTTCGTTTCTTCACTGCAATC ATTTTTTTTTTTGGTCGTTCATATGTTCTCCTTCCTAGTACAATTTGTGCATTTT CATTGTTTTCTAGAATTCCTTCATTTTGGCTTTATTCTAGCACCATTTTTCACA TTTTCTGTCCCCACACACACACAAAATACATTGCCAATTATATATTTTTAATTA CAAATTCCCCCAAAAATCCGTTGTTGTTTTCTTTTCTCCGGTGATTGACACTT TTTATTTCATTTTTTGTAAAATATTACACATTCGCCGGTTTTTCCATCGAAACT GACTAACGTGTAATACTAATAATCGTATTCAAAATGTTTTTTTTTTCAAATTCCT GCCACATTCTTCTTCATCTGTATGCGTTTCAAACGCCCCATCCATCCACCCG GTTGCCCTcgataacatggtcatagctgtttcctgtgtgaaattg

[a]Overhangs are in lowercase, the part annealing to the template in uppercase. If not stated otherwise, the PCR products were amplified from *C. elegans* genomic DNA.
[b]Overhangs for Gibson assembly reactions are in lowercase.
[c]Assembly with EcoRV digested pDONR L4-R1 *lin-29A* promoter plasmid and two PCR products. PCR amplification from pDONR L4-R1 *lin-29A* promoter plasmid.
[d]Assembly with EcoRV digested pDONR L4-R1 *lin-29A* promoter plasmid and one gBlock.
[e]Assembly with XbaI digested pBMF2.7 plasmid and two PCR products. PCR amplification with fwd(1) and rev(1) from pDONR L4-R1 *lin-29A* promoter plasmid, with fwd(2) and rev(2) from pBMF2.7 plasmid.
[f]Assembly with XbaI digested pBMF2.7 plasmid, one gBlock and one PCR product. PCR amplification with fwd(2) and rev(2) from pBMF2.7 plasmid.
[g]PCR amplification of *lin-41* CDS (genomic sequence) with fwd(1) and rev(1), PCR amplification of flag::gfp with fwd(2) and rev(2). Annealing sequences for fusion PCR with fwd(1) and rev(2) are indicated in bold letters.
[h]Assembly with AscI digested modified pCM5.37 plasmid (G. Brancati and H. Großhans) and one gBlock.
[i]Assembly with ClaI and SpeI digested pDD282 plasmid and two gBlocks.

2.3 LIN41 may recognize an RNA structure rather than a linear motif

At least in the *C. elegans* soma, the RNA-binding protein LIN41 seems to bind very specifically to only a few target mRNAs ((Aeschimann et al., 2017), see also section 2.4 below). We therefore wondered how LIN41 would achieve such specificity. For another TRIM-NHL protein, *Drosophila* Brat, multiple expectation and maximization for motif elicitation (MEME) analysis (Bailey and Elkan, 1994) on its identified targets, in combination with RNAcompete experiments (Ray et al., 2009), suggested a short motif ((U/A)UGUUA) as its recognition site on target mRNAs (Loedige et al., 2015). Indeed, a crystal structure revealed that the Brat NHL domain can contact this motif in a highly sequence-specific manner (Loedige et al., 2015).

2.3.1 LIN41 is unlikely to bind to a short linear motif

Loedige *et al*. also performed RNAcompete experiments on other TRIM-NHL proteins including *C. elegans*, fly and human LIN41 homologs, but the identified short RNA sequence motifs were less clearly defined compared to the motif of Brat. When we searched for the proposed LIN41 motifs, we found them to be rather depleted than enriched among our four identified LIN41 targets in *C. elegans* (data not shown). Furthermore, MEME analysis with the 3'UTRs of *mab-10*, *mab-3* and *dmd-3* and the 5'UTR of *lin-29A* did not yield any obvious, enriched sequence motifs (data not shown). However, this may be expected from analysis with such a small set of targets. As described in the publication of section 2.2, we took a different strategy to get closer at the sequence LIN41 binds to by trying to narrow down the element that confers LIN41-mediated repression in the 5'UTR exon of *lin-29A*. To do so, we constructed reporter transgenes consisting of GFP(PEST)::H2B, driven from the ubiquitously expressing *dpy-30* promoter, and fused to 100-nt long, overlapping stretches of the *lin-29A* 5'UTR. To our surprise, whereas the whole *lin-29A* 5'UTR resulted in efficient repression, none of the three tested fragments were sufficient to confer LIN41-mediated repression on the reporter (Figure S6 in (Aeschimann et al., 2017)). We conclude that LIN41-mediated repression is unlikely to result from the binding of a LIN41 protein to a short linear stretch of RNA on the *lin-29A* 5'UTR. Rather, LIN41 may need multiple binding sites for target repression. As another explanation for these results, the binding motif for LIN41 may not be exposed for binding in the tested sequence contexts.

2.3.2 LIN41 may bind to an RNA structure

In an alternative scenario, LIN41 recognizes an RNA structure, rather than a short linear motif like Brat. Hence, we analyzed the predicted RNA secondary structures of the *lin-29A* 5'UTR exon using RNAfold (Lorenz et al., 2011) (Figure 4A). In this prediction, two out of the several structured regions caught our attention: First, a rather long stretch of double-stranded RNA (dsRNA), interrupted by a bulge, that was predicted to form between the 5' most and 3' most nucleotides of the sequence (Figure 4A, colored in blue). Second, two very similar stem loops predicted to occur approximately in the middle of the sequence (Figure 4A, colored in red, and Figure 4C). As the long dsRNA region could not have formed in any of our tested fragments, we additionally constructed a reporter lacking just the two similar stem loops, which should allow formation of the dsRNA region (Figure 4B, deletion of the nucleotides colored in red in Figure 4A). However, this reporter was also unable to induce LIN41-mediated repression, suggesting that the long double-stranded RNA region is at least not sufficient to confer repression by LIN41. In addition, the two stem loops could be involved in repression, but are probably also not sufficient, as the middle fragment of the lin-29A 5'UTR exon (nucleotides 52-155), which contains the nucleotides of both stem loop structures, was unable to repress the reporter (Figure S6 in (Aeschimann et al., 2017)). Nevertheless, we searched for similar stem loops in the 3'UTRs of the other LIN41 targets and indeed found some within the predicted structures (Examples are shown in Figure 4C). This could be mere coincidence, but it would be interesting to test if deletion of these predicted structures within the long 3'UTRs would also affect silencing by LIN41.

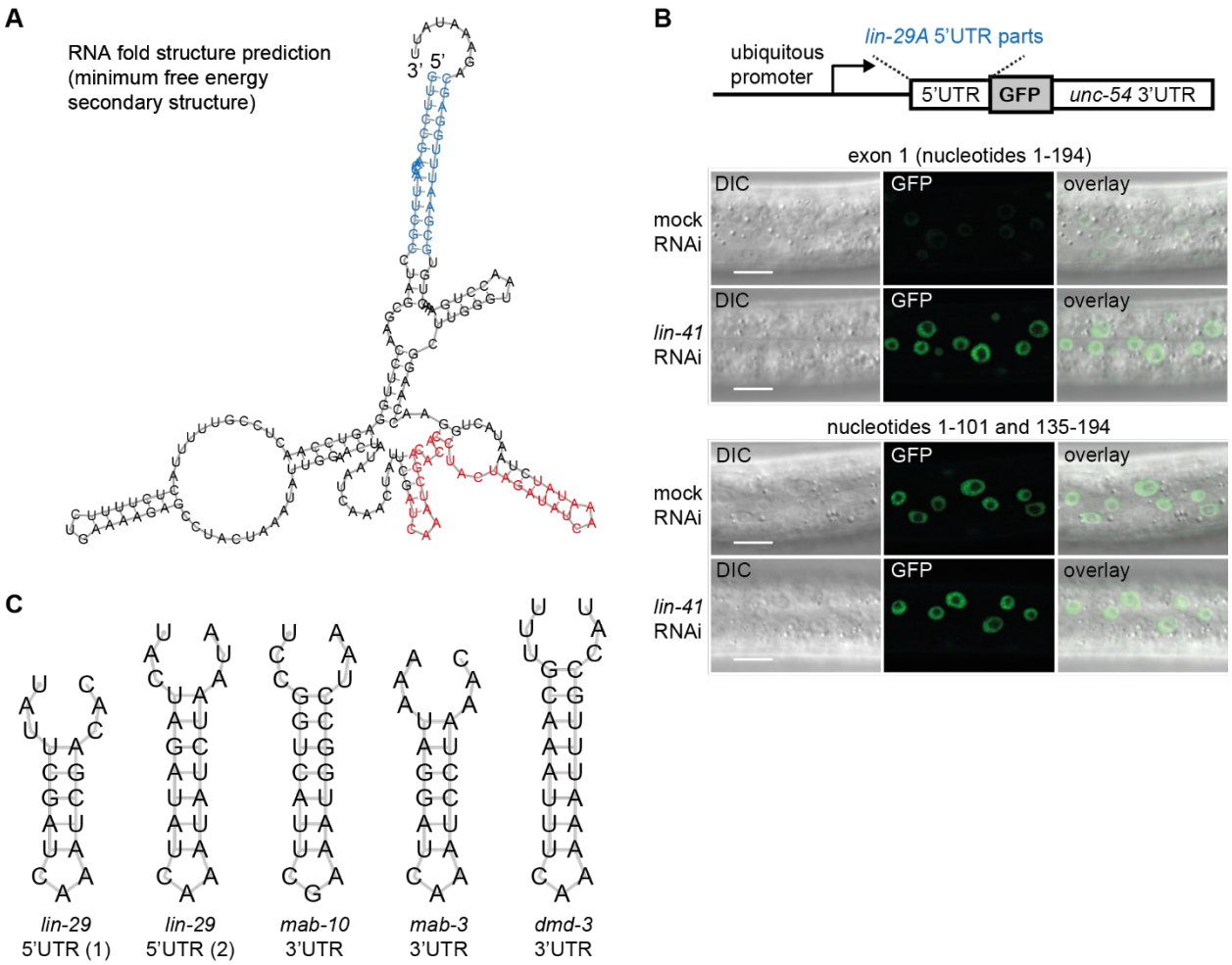


Figure 4. The sequence or structure requirements for RNA-binding by LIN41

(A) Prediction of the secondary structures in the *lin-29A* 5'UTR exon using RNAfold (Lorenz et al., 2011).

(B) *In vivo* confocal imaging of nuclear GFP(PEST)::H2B driven from the *dpy-30* promoter and fused to the *unc-54* 3'UTR and the *lin-29A* 5'UTR exon or the indicated part thereof. Imaged were epidermal nuclei of L3 larval stage worms, grown in the presence of *lin-41* RNAi or mock RNAi bacteria. Shown are images with differential interference contrast (DIC), GFP (identical settings for both RNAi conditions) and the overlay of the two. Scale bars: 10 µm. The upper image was also used for Figure S6 in (Aeschimann et al., 2017).

(C) Similar stem-loop structures found in predicted secondary structures of the LIN41 target UTRs.

2.4 LIN41 specifically binds to only a few somatic mRNAs

Through analysis of ribosome profiling time courses, we identified four somatic LIN41 targets, *mab-3*, *dmd-3*, *lin-29A* and *mab-10* (Aeschimann et al., 2017). We therefore wondered if LIN41 regulates only four somatic targets or if we had been unable to identify more targets using the ribosome profiling assay. To identify LIN41 targets in another, unbiased manner, we performed RNA co-immunoprecipitations coupled to RNA sequencing (RIP-seq) with a semi-synchronous population of animals in the L3 and L4 larval stages, during which LIN41 regulates the four identified mRNA targets. Using an anti-FLAG antibody, we applied RIP-seq to worms with transgenic expression of FLAG::GFP::LIN41 in the *lin-41(n2914)* mutant background and to wild-type worms expressing FLAG::GFP::SART-3 (Rüegger et al., 2015) as a control. Because two of the four previously identified LIN41 targets, *mab-3* and *dmd-3*, are mainly expressed in male worms (Mason et al., 2008), we additionally performed RIP-seq on worm populations in the *him-5(e1490)* genetic background, enhancing the frequency of males in the population from less than 1 to about 35 percent (Meneely et al., 2012).

2.4.1 LIN41 may regulate its expression by binding to its own mRNA

For each LIN41 RNA-IP, we identified RNAs enriched at least 4-fold (2-fold in $\log_2$ scale) compared to the SART-3 IP, and selected those meeting this criterion in at least two out of the three biological replicates (Figure 5A). For the IPs from worm populations with the normal male frequency, this selection included only four enriched mRNAs: *lin-29*, *mab-10*, *mab-3* and *lin-41*. This result suggested that indeed there seem to be only few somatic mRNA targets of LIN41 and that we had already identified almost all of them in the ribosome profiling experiment. The only previously known somatic LIN41 target that was not enriched in these IPs was *dmd-3*. Interestingly, with *lin-41*, we identified a new target, suggesting that LIN41 binds to its own mRNA. Strikingly, the enrichment of *lin-41* was even higher than that of *lin-29* or *mab-10*, in both the RIP-seq data (Figure 5A) and the RT-qPCR analysis (Figure 5B). Previously, LIN41 has been proposed to autoregulate its activity (Del Rio-Albrechtsen et al., 2006), but rather on the basis of its function as an E3 ubiquitin ligase. Consistent with LIN41 regulating its own expression, we observed a lower GFP intensity of the FLAG::GFP::LIN41 transgene in wild-type compared to the *lin-41(n2914)* mutant background (data not shown). This difference was only apparent in the germline of adult hermaphrodites, as the cytoplasmic LIN41 signal was hard to detect in any other tissue and any other developmental stage. Preliminary results (data not shown) suggest that reporter lines containing the *lin-41* 5'UTR and/or *lin-41* 3'UTR seem not to be repressed by LIN41 in L3 stage animals, as GFP levels did not change upon crossing them into a reduction-of-function and temperature-sensitive allele *lin-41(tn1487ts)* (Spike et al., 2014). However, we

cannot exclude that expression of these reporters would be changed in conditions with complete absence or strong overexpression of LIN41, for example in *lin-41(n2914)* or *lin-41(xe8)* mutants, respectively. For now, we conclude that LIN41 likely regulates its own expression, possibly by binding to its own mRNA, but future experiments are needed to elucidate the mechanisms of this autoregulatory loop (see also Discussion).

2.4.2 LIN41 targets a small set of mRNAs in the soma

For the RNA-IPs from worm populations with the enhanced male frequency, we selected six mRNAs as specifically enriched in LIN41 IPs. Next to *lin-29*, *mab-10*, *mab-3* and *lin-41,* the selection included *dmd-3* and the unannotated gene *F18C5.10* (Figure 5A). A higher percentage of males in the population, clearly increasing the read numbers for *dmd-3* (data not shown), thus enabled us to detect its binding to LIN41. In RNA-IPs followed by RT-qPCR analysis in the publication in section 2.2, we found that both *mab-3* and *dmd-3* were clearly less enriched in LIN41 IPs compared to *lin-29* and *mab-10*. However, as both *mab-3* and *dmd-3* are only expressed at low levels in hermaphrodites (Mason et al., 2008), this could simply reflect a technical issue. We therefore quantified the enrichments of *mab-3* and *dmd-3* by RT-qPCR in IPs performed from *him-5* mutant worm populations (Figure 5C). These enrichments were still modest, especially when compared to those of *lin-29* and *mab-10*, suggesting that *mab-3* and *dmd-3* could be bound by LIN41 with lower affinity and/or in fewer cells. The additional mRNA enriched in the LIN41 IPs in *him-5* background (Figure 5A), *F18C5.10*, was not changed in the ribosome profiling experiment and we did not pursue it further. We conclude that during larval development, LIN41 likely regulates only the four previously identified targets, *lin-29A*, *mab*-10, *mab-3* and *dmd-3*, plus its own mRNA.

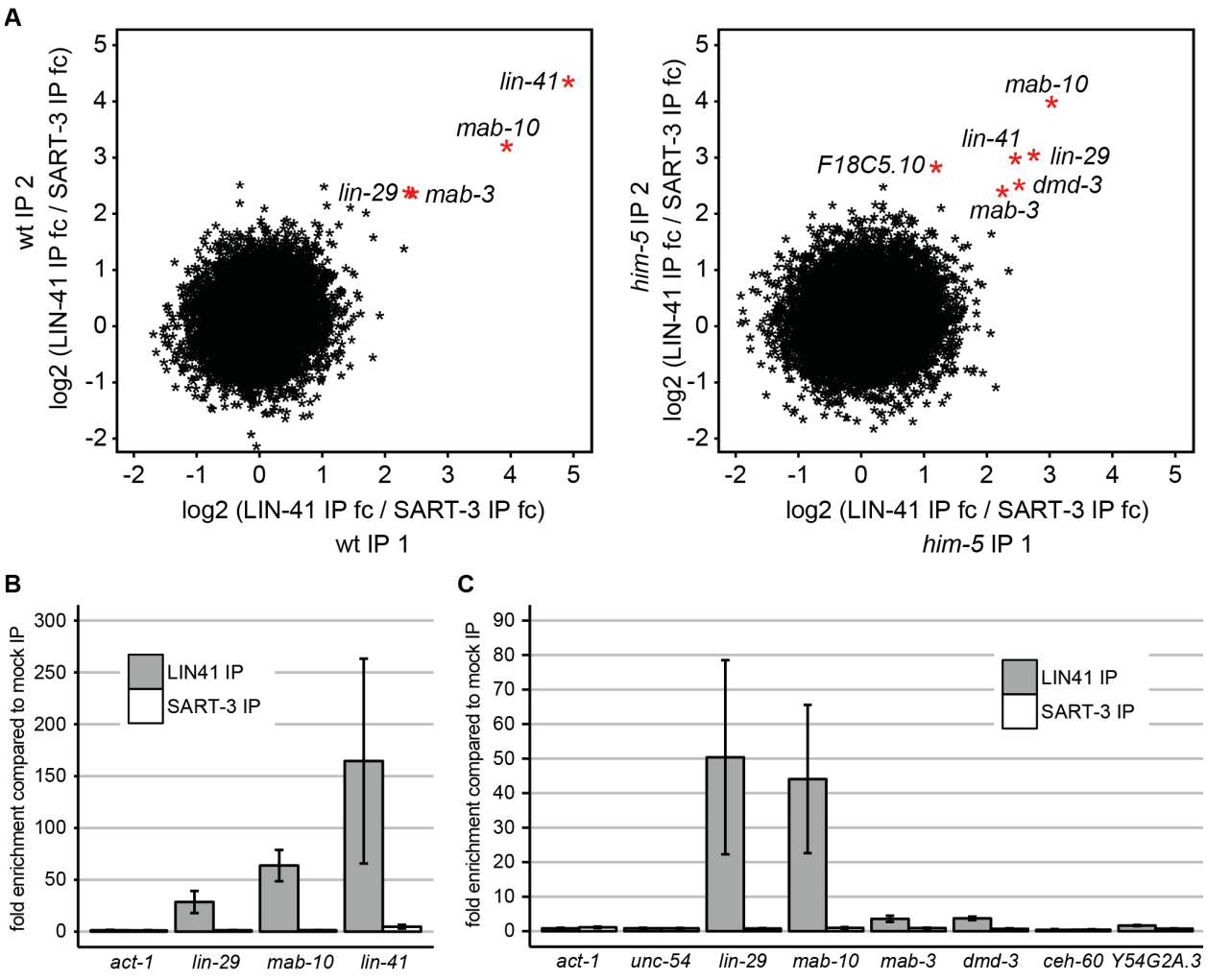


Figure 5. Identification of LIN41 mRNA targets by RIP-seq

(A) Fold changes in anti-FLAG RNA co-immunoprecipitation coupled to RNA sequencing (RIP-seq) of FLAG::GFP::LIN41 compared to FLAG::GFP::SART-3 (Rüegger et al., 2015) as a control. Depicted are a scatter plot comparing the fold changes in two (out of three) biological replicates, with experiments using populations with low and high frequency of males shown in the left and right panel, respectively. Marked in red are mRNAs with fold changes of at least 2 (log_2 scale) in at least two of the three biological replicates.

(B) RT-qPCR analysis to test for co-immunoprecipitation (co-IP) of candidate LIN41 target mRNAs with FLAG::GFP::LIN41 or FLAG::GFP::SART-3 in semi-synchronous L3/L4 stage worm populations. Fold enrichments were calculated relative to anti-FLAG IP in non-transgenic, wild-type animals (mock IP). n = 4 biological replicates, data as mean ± s.e.m.

(C) RT-qPCR analysis as in (B), but with semi-synchronous L3/L4 stage worm populations of increased male frequency (*him-5(e1490)* genetic background). n = 3 biological replicates, data as mean ± s.e.m.

2.5 LIN41 controls development of female and male sexual organs through different targets

Next to the defects in epidermal development, two additional somatic phenotypes have been described for *lin-41* gain-of-function (gf) mutants, a defect in the development of the vulva (Ecsedi et al., 2015) and a defect in the development of the male tail tip (Del Rio-Albrechtsen et al., 2006). Contrary to the phenotype in seam cells, both defects are observed on the level of morphogenesis of the sexual organs without any reported changes in cell proliferation (Del Rio-Albrechtsen et al., 2006; Ecsedi et al., 2015). LIN41 thus has a role in development of both, female and male sexual organs, but how it achieves this function and through which targets it acts has been unknown. The combination of our genome-wide ribosome profiling time course and RIP-seq experiments ((Aeschimann et al., 2017) and section 2.4) suggested that LIN41 only regulates five somatic targets, *lin-29A, mab-10*, *mab-3*, *dmd-3* and *lin-41*. Assuming all observed *lin-41* mutant phenotypes result from dysregulation of LIN41 targets other than its own mRNA, we hypothesized that this small set of four targets could be responsible for all observed somatic *lin-41* mutant phenotypes.

2.5.1 Depletion of LIN-29 and MAB-10 can cause vulval bursting

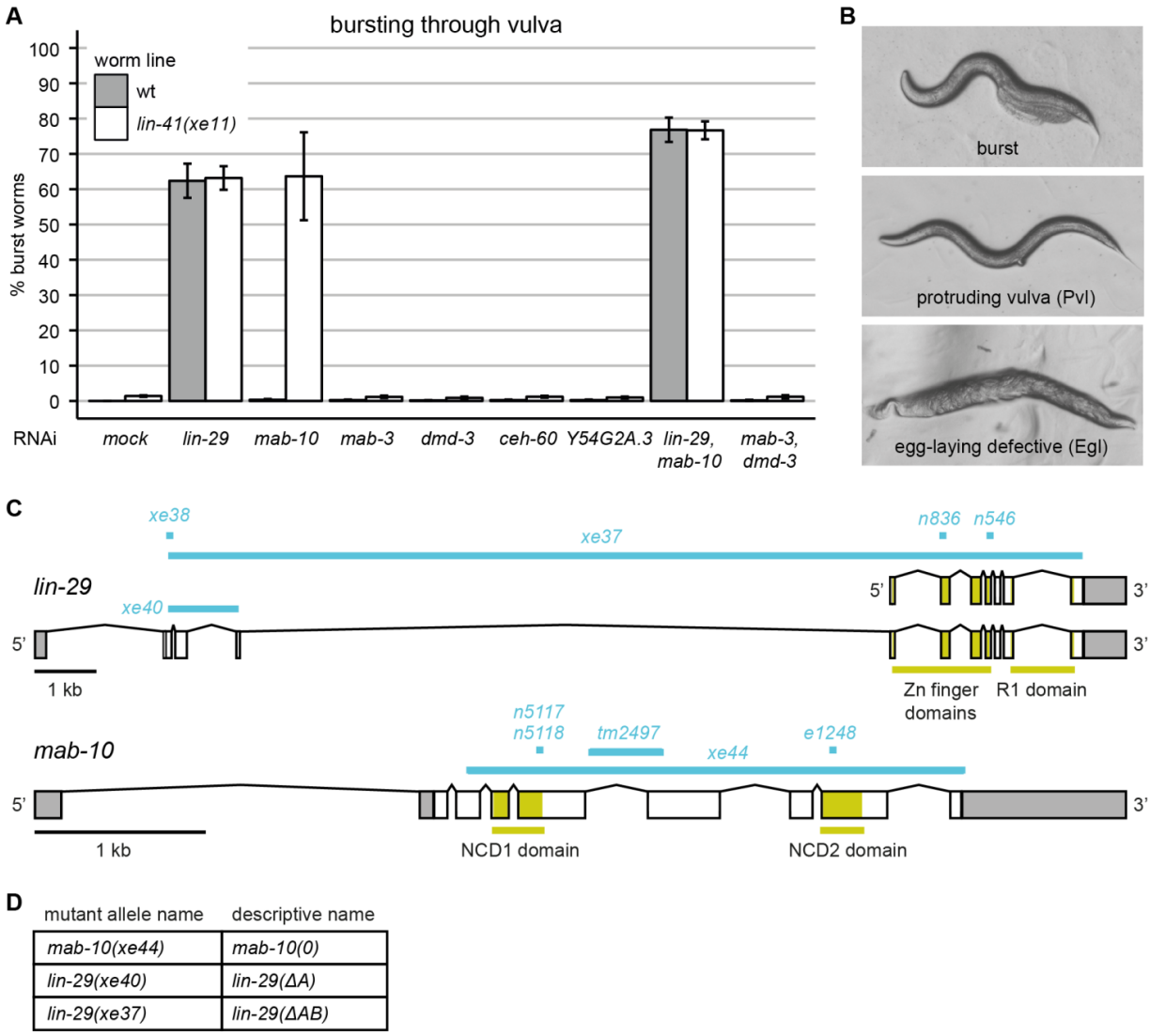
The defect in the development of the vulva is the cause of lethality in *lin-41(xe8)* gf mutant animals (Ecsedi et al., 2015). These mutants lack the *let-7* complementary sites in the *lin-41* 3'UTR and perfectly phenocopy the *let-7* mutant phenotype, as they burst through their vulva shortly after the L4-to-adult molt (Ecsedi et al., 2015). As point mutations in the *lin-41* 3'UTR, re-establishing binding to the mutant version of *let-7*, rescue the inviability of *let-7(n2853)* mutant animals, LIN41 was found to be the sole relevant target for the function of *let-7* in vulval development (Ecsedi et al., 2015). Whereas elevated LIN41 levels can fully explain why *let-7* mutant worms are inviable, it is in turn unknown why these elevated LIN41 levels do not allow for survival into adulthood. As an alternative to LIN41 as a post-transcriptional gene regulator, it was proposed that the LIN41 protein could directly contribute to structural integrity of the vulva (Ecsedi et al., 2015), as its fly homolog dappled/wech has been shown to link integrins to the cytoskeleton (Loer et al., 2008). Consistent with this idea, vulval bursting has not been reported for mutants of *lin-29*, *mab-10*, *mab-3* or *dmd-3.* Taken together, the vulva bursting phenotype, elicited by sustained LIN41 protein production in late larval stages, has remained a conundrum.

We examined whether repression of any of the identified LIN41 targets could contribute to the defect in vulval development caused by elevated LIN41 levels. We hypothesized that it might be a combination of LIN41 targets, rather than a single LIN41 target, that would cause vulval rupturing if they are not properly expressed at the larval-to-adult transition. We therefore scored for burst animals after feeding them with

bacteria inducing RNAi against the individual LIN41 targets, as well as combinations of *lin-29* and *mab-10* or *mab-3* and *dmd-3*, because each of these gene pairs had been shown to act in the same developmental pathway (Harris and Horvitz, 2011; Mason et al., 2008). In addition, we tested the depletion of *ceh-60* and *Y54G2A.3*. These genes, although not enriched in LIN41 RNA-IPs, were changed due to and immediately after mis-expression of LIN41 in a time course experiment (Aeschimann et al., 2017). Moreover, we screened for bursting in a sensitized genetic background using *lin-41(xe11)* gf mutant animals, which do not show the bursting phenotype (Ecsedi et al., 2015), but express elevated LIN41 levels and therefore decreased LIN41 target levels (Aeschimann et al., 2017). We reasoned that performing the experiment with *lin-41(xe11)* animals would increase our chances to identify LIN41 targets involved in bursting, even if they were redundant with other targets for their function in vulval development.

Depletion of individual LIN41 targets in wild-type worms did not cause any bursting, with the prominent exception of worms depleted for LIN-29, where more than 60 % of animals died by bursting through their vulva as young adults (Figure 6A). This was a surprising result, as vulval bursting had not been reported for *lin-29(0)* mutant animals. In the sensitized *lin-41(xe11)* mutant background, depletion of LIN-29 caused a similar percentage of animals to burst, but additionally, depletion of MAB-10 caused bursting (Figure 6A). Consequently, double depletions of LIN-29 and MAB-10 caused vulval bursting in both wild-type and *lin-41(xe11)* animals (Figure 6A). As MAB-10 is a co-factor of the transcription factor LIN-29 (Harris and Horvitz, 2011), these results suggested that a change in activity of LIN-29, acting with its co-factor MAB-10, can cause lethality through vulval bursting.

Whereas vulval development has no reported defects in *mab-10(0)* mutants, it is abnormal in *lin-29(0)* mutant animals (Ambros and Horvitz, 1984; Bettinger et al., 1997; Inoue et al., 2005; Newman et al., 2000). First, *lin-29(0)* mutants are egg-laying defective (Egl), i.e. unable to lay eggs through the vulva, leading to embryonic development and hatching inside the mother animal (Figure 6B) (Ambros and Horvitz, 1984; Bettinger et al., 1997). Second, they have a protruding vulva (Pvl), i.e. an abnormal ventral protrusion where the vulva forms (Figure 6B) (Ambros and Horvitz, 1984; Bettinger et al., 1997). Third, they exhibit a defect in the vulva to uterus connection and in expression of cell-specific markers in uterine and vulval cells (Inoue et al., 2005; Newman et al., 2000). However, because all these defects were not observed in *let-7* mutants, whereas bursting was not observed in *lin-29(0)* animals, the *let-7* and *lin-29* mutant phenotypes seemed to be completely different, which has led to the conclusion that LIN-29 is not a key effector for the function of *let-7* in vulval development (Ecsedi et al., 2015).



mutant allele name	descriptive name
mab-10(xe44)	*mab-10(0)*
lin-29(xe40)	*lin-29(ΔA)*
lin-29(xe37)	*lin-29(ΔAB)*

Figure 6. Depletion of LIN-29 and MAB-10 causes vulval bursting.
(A) Quantification of burst wild-type or *lin-41(xe11)* mutant worms depleted of the indicated mRNAs through RNAi. n = 3 biological replicates, data as mean ± s.e.m. At least 400 worms were counted for each condition.
(B) Example pictures to illustrate the burst, Pvl and Egl phenotypes. Shown is a burst *lin-29(xe40); mab-10(xe44),* a Pvl *lin-29(xe40)* and an Egl *lin-29(xe37); mab-10(xe44)* mutant animal.
(C) Illustration of different mutant alleles (in blue) of *lin-29* and *mab-10*.
(D) Table of the descriptive mutant allele names used in this study.

Because the results of the RNAi experiment suggested a role for LIN-29 in the vulval bursting phenotype, we tested if, although not reported, vulval bursting can occur in *lin-29(0)* mutants. Surprisingly, when comparing two of the published *lin-29(0)* mutants, *lin-29(n546)* and *lin-29(n836)* (Ambros and Horvitz, 1984), we noticed obvious phenotypic differences between them. Whereas their phenotypes in the epidermis may be similar (Ambros and Horvitz, 1984), they were clearly different in their vulval defects

and survival rates, with *lin-29(n546)* mutants often dying by bursting through their vulva and *lin-29(n836)* mutants only rarely bursting and usually surviving into adulthood (phenotypes not quantified, data not shown). Intuitively, this would suggest that *lin-29(n546)*, but not *lin-29(n836)*, is a *null* allele, and that *lin-29(0)* mutants show the bursting phenotype. However, this conclusion is challenged by analysis of the mutants at the DNA level. The *lin-29(n546)* allele is a point mutation leading to a premature stop codon in the *lin-29* ORF, just downstream of the zinc finger domains (Figure 6C). When in a previous study, production of the C-terminally truncated protein was stabilized by downregulation of the nonsense-mediated mRNA decay (NMD) pathway, *lin-29(n546)* mutant animals were rescued for almost all phenotypes (Harris and Horvitz, 2011). As mRNA degradation through the NMD pathway is usually not 100 % efficient (Schweingruber et al., 2013), this would suggest that *lin-29(n546)* is not a *null* mutant. By contrast, *lin-29(n836)* is a short deletion of 73 bp, starting at the first nucleotide of exon 6 (Figure 6C), presumably resulting in a frame-shift and premature stop codon within the zinc finger domains and thus more likely to be a *null* mutant. Due to these uncertainties with the *lin-29(n546)* and *lin-29(n836)* alleles, we created a new allele *lin-29(xe37)* that was undoubtedly *null* by deleting almost the entire coding region of the long isoform *lin-29A* (Figure 6C, see Supplemental Methods for details). Its phenotype, as described in more detail below, resembled the *lin-29(n836)* allele and not the phenotypically stronger *lin-29(n546)* allele. From the analysis of *lin-29* mutants and the RNAi experiment, we conclude that a partial rather than a complete loss of LIN-29 activity could be responsible for bursting through the vulva.

2.5.2 LIN41 regulates LIN-29A and MAB-10 to ensure proper vulval development

We previously found that LIN41 targets only one LIN-29 isoform, LIN-29A, without targeting the LIN-29B isoform (Aeschimann et al., 2017). As the *lin-29(0)* mutants lack expression of both isoforms, we were interested whether isoform-specific dysregulation could contribute to the lethal phenotype. We therefore disrupted specifically *lin-29A*, without affecting expression of *lin-29B*, using CRISPR-Cas9. The resulting *lin-29(xe40)* allele deletes coding exons two through four of *lin-29* while causing a frame-shift in the remaining *lin-29A* open reading frame, rendering it *lin-29A(0)* (Figure 6C). We then quantified bursting of young adult worms in *lin-29(xe40)* mutant as well as in wild-type, *lin-29(xe37), let-7(n2853)* and *lin-41(xe8)* genetic backgrounds. In order to use more descriptive names for the different *lin-29* mutants, we refer to the *lin-29(xe37)* null allele as *lin-29(ΔAB)* and to the *lin-29A* specific allele *lin-29(xe40)* as *lin-29(ΔA)* (Figure 6D). While at the time point of analysis, about 90 % of the *let-7(n2853)* and *lin-41(xe8)* worms had burst through their vulva, only very few *lin-29(ΔA)* or *lin-29(ΔAB)* worms had burst (Figure 7A, see also Table S1 for exact percentages). However, we confirmed the previously reported Pvl and Egl phenotypes of *lin-29(0)* mutant worms (Ambros and Horvitz, 1984; Bettinger et al., 1997) and, by quantifying these

phenotypes, found that almost 100 % of *lin-29(ΔAB)* worms had a protruding vulva as young adults and were egg-laying defective as older adults (Figure 7, see also Tables S2 and S3 for exact percentages). Specific mutation of the LIN-29A isoform in *lin-29(ΔA)* mutants also resulted in Pvl and Egl phenotypes, although with slightly lower penetrance (Figure 7). Although we did not quantify the phenotypes, we confirmed that specific loss of LIN-29A results in Pvl and Egl, but not bursting phenotypes, with a second and independently obtained *lin-29A* specific mutant allele, *lin-29(xe38)* (Figure 6C, see Supplemental Methods for details). As LIN41 downregulates not only LIN-29A, but also its co-factor MAB-10, we next tested whether *mab-10*; *lin-29A* double mutants would die by bursting through the vulva. For this purpose, we generated a *mab-10(xe44)* deletion strain lacking most coding exons including the functionally important NCD domains (Harris and Horvitz, 2011) (Figure 6C). We then crossed the *lin-29(ΔA)* allele with the *mab-10(xe44)* null allele, referred to as *mab-10(0)* (Figure 6D). Indeed, this double depletion caused vulval bursting in about 30 % of worms, while the surviving worms showed fully penetrant Pvl and Egl phenotypes (Figure 7). Complete loss of LIN-29 and MAB-10 in *lin-29(ΔAB); mab-10(0)* double mutants led to bursting with much lower frequency, but to penetrant Pvl and Egl phenotypes, as observed in *lin-29(ΔAB)* worms (Figure 7). We verified the enhanced frequency of burst worms in *lin-29(ΔA); mab-10(0)* mutants compared to *lin-29(ΔAB); mab-10(0)* mutants with different combinations of alternative *lin-29(0), lin-29A* and *mab-10* mutants (phenotypes not quantified). Importantly, the *mab-10(0)* mutation alone did not cause any bursting, nor any other obvious vulva phenotypes, except for an Egl phenotype with low penetrance (Figure 7). In summary, the vast majority of worms fully depleted of LIN-29 activity, in *lin-29(ΔAB)* or *lin-29(ΔAB); mab-10(0)* mutant backgrounds, survive into later adulthood, while about 30 % of worms depleted of only MAB-10 and LIN-29A, but not LIN-29B, burst shortly after the larval-to-adult transition. Notably, we observed bursting in about 50 % of *lin-29(ΔA); mab-10(0)* mutant animals after growing them at lower temperature (15 °C, see Table S4 for percentages). However, although we can observe the bursting phenotype in *lin-29(ΔA); mab-10(0)* mutant animals, this cannot explain the almost complete penetrance of the phenotype in *let-7(n2853)* and *lin-41(xe8)* mutant animals (Figure 7A). In turn, this would suggest that this complete penetrance could be due to tissue-specific mis-expression of LIN-29A and MAB-10, mis-expression of additional LIN41 targets, the role of LIN41 as a structural protein or a combination thereof (see Discussion).

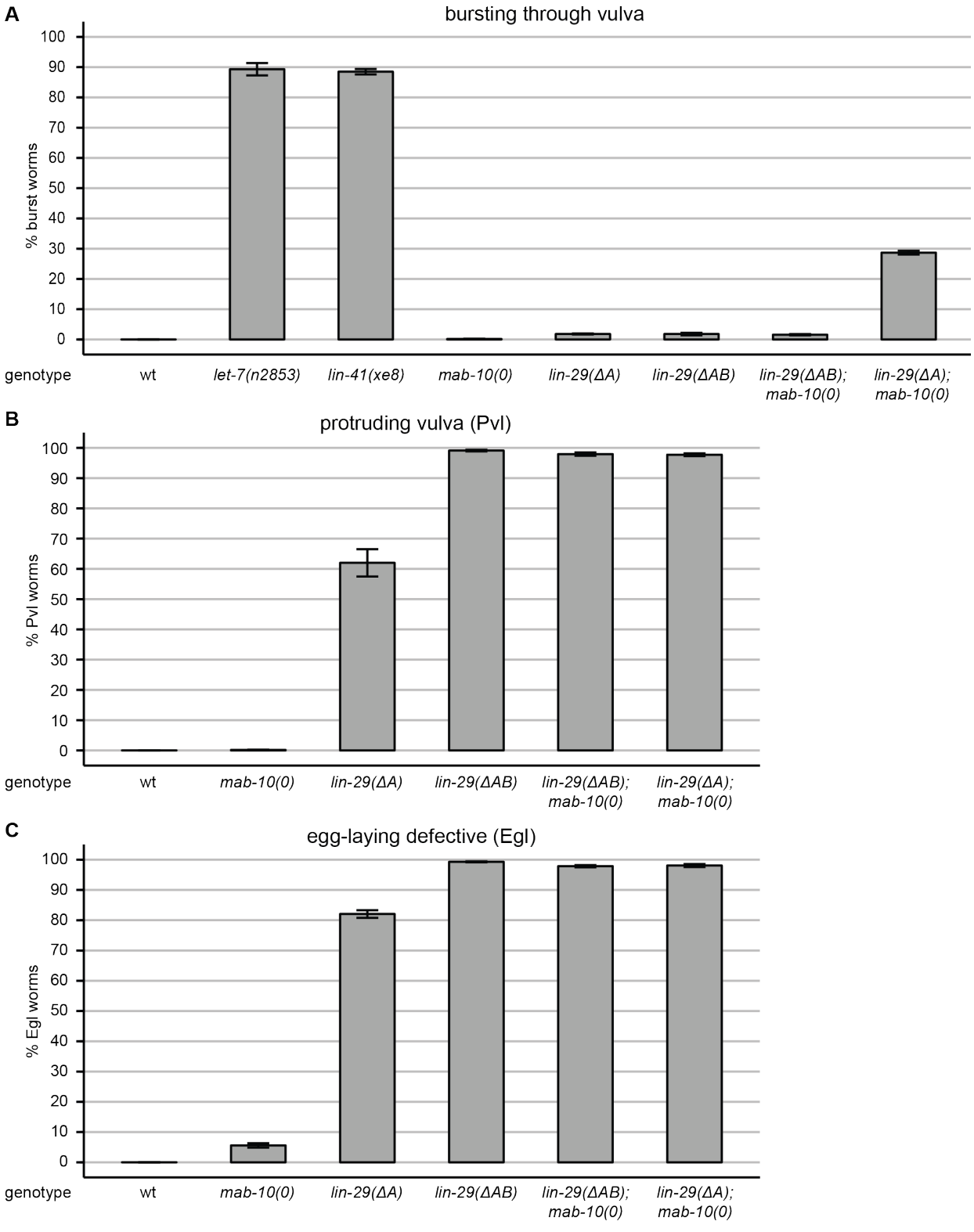


Figure 7. The bursting, Pvl and Egl phenotypes in worms of different *lin-29* and/or *mab-10* mutant backgrounds
(A-C) Quantification of burst (A), Pvl (B) or Egl (C) worms of the indicated genotypes. n = 3 biological replicates, data as mean ± s.e.m. At least 400 worms were counted for each condition.

2.5.3 LIN41 regulates MAB-3 and DMD-3 to ensure male tail retraction at the right time

In the early larval stages L1 through L3, the tails of *C. elegans* hermaphrodites and males are both pointed and morphologically very similar. While hermaphrodites retain the pointed tail also in the L4 and adult stage, a dramatic remodelling event re-shapes the male tail tip during the L4 stage, resulting in a blunt-ended, rounded tail tip. This morphogenesis is necessary for successful mating with the hermaphrodite and is brought about by anterior retraction of the most posterior epidermal cells of the tail. Defects in this retraction result in abnormally pointed tail tips, referred to as "leptoderan" (Lep), a phenotype observed in males harbouring the gf alleles *lin-41(bx37)* or *lin-41(bx42)* (Del Rio-Albrechtsen et al., 2006). The observed cause of this defect is a delay in the onset of cell retraction, resulting in only partially retracted tail tips. By contrast, the tail tips of males with the reduction-of-function (rf) allele *lin-41(ma104)* retract too early, resulting in over-retracted tips (Del Rio-Albrechtsen et al., 2006).

As for the defect in vulval development, it was unclear whether any of the involved LIN41 targets could be responsible for the defect in male tail development or whether, for example, LIN41 rather plays a structural role in integrin to cytoskeleton attachments during male tail retraction. Interestingly, abnormalities in male tail tips have been reported for mutants of *lin-29*, *mab-3* and *dmd-3* (Euling et al., 1999; Mason et al., 2008). While *lin-29(0)* mutant males are mating defective due to problems in spicule formation and ray extension, they exhibit apparently normal cell retraction (Euling et al., 1999). By contrast, *dmd-3* mutant males have a Lep phenotype with partially unretracted tail tips, whereas *mab-3; dmd-3* double mutant males have completely unretracted male tails (Mason et al., 2008). Out of these three phenotypes, the rather mild *lin-41(gf)* phenotype seems most similar to that of *dmd-3* single mutant males, suggesting it could be caused by specific dysregulation of *dmd-3* in male tails. Indeed, DMD-3 expression in male tails has previously been proposed to be regulated by LIN41, through an unknown mechanism involving the DMD-3 promoter (Mason et al., 2008). Assuming LIN41 acts as an RBP in this pathway, this could suggest an indirect mechanism, whereby a transcription factor downstream of *lin-41* would regulate *dmd-3* transcription.

Since we previously found that the 3'UTRs of both *mab-3* and *dmd-3* harbor a binding site for the LIN41 protein (Aeschimann et al., 2017), we wanted to test the alternative possibility that LIN41 directly regulates *mab-3* and/or *dmd-3* mRNAs in male tails. Thus, we created reporter lines to express GFP(PEST)::H2B from *mab-3* and *dmd-3* promoters, i.e. the 4 kb region upstream of their start codons, and fused to their 3'UTR regions to include the identified LIN41 target sites. Consistent with a previous report using translational reporters (Mason et al., 2008), our reporters were only very weakly expressed in hermaphrodites, but accumulated a strong GFP signal in L4 stage males (data not shown). We next examined GFP signals in early L3 stage males, when LIN41 levels are still high. Consistent with potential

inhibition through LIN41, GFP expression in most tissues was almost undetectable for both reporters, with the exception of the pharynx region in the *mab-3* reporter and the distal tip cell in the *dmd-3* reporter (data not shown). However, upon depletion of LIN41 by RNAi, we detected strong GFP signals in many tissues for both reporters, including the epidermal cells of the tail region (Figure 8), suggesting that LIN41 represses the expression of *mab-3* and *dmd-3* in these cells. In addition to the tail region, both reporters were clearly de-repressed in epidermal hyp7, while the *mab-3* reporter was additionally de-repressed in cells of the gut and the somatic gonad (data not shown).

To test if LIN41-mediated regulation of *mab-3* and *dmd-3* expression was mediated by their 3'UTRs, we exchanged the reporter 3'UTRs with the unregulated *unc-54* 3'UTR. For both, *mab-3* and *dmd-3* reporters, this resulted in strong GFP expression in all the cells we previously found to show LIN41-dependent regulation, including the epidermal cells of the tail region (Figure 8). Consistent with LIN41 repressing *mab-3* and *dmd-3* only via their 3'UTRs, we did not see any change in GFP levels upon LIN41 depletion (Figure 8). We conclude that, in males, LIN41 regulates *mab-3* and *dmd-3* through a post-transcriptional mechanism by binding to their 3'UTRs. Expression of both reporters in hermaphrodites was still weak upon depletion of LIN41, suggesting that additional mechanisms regulate the male-specific expression of *mab-3* and *dmd-3*, likely at the transcriptional level (Mason et al., 2008).

The extensive repression by LIN41 in the male tail region (Figure 8) suggested that *mab-3* and *dmd-3* expression is completely silenced in the presence of LIN41. Consequently, a failure in LIN41 downregulation in *let-7(lf)* and *lin-41(gf)* mutants should result in almost complete absence of MAB-3 and DMD-3 proteins in the tail epidermis of L4 stage males. In turn, this should cause completely unretracted male tails, as seen for *mab-3; dmd-3* double mutant males. However, both *let-7(lf)* and *lin-41(gf)* males were reported to only display a partial defect in tail retraction (Del Rio-Albrechtsen et al., 2006). At first glance, these results could suggest that silencing of *mab-3* and *dmd-3* by LIN41 is only partial, or that silencing of additional targets like *mab-10* or *lin-29A* would partially suppress the complete failure of cell retraction. A closer look at the published observations reveals that the described *let-7(lf)* and *lin-41(gf)* phenotypes were those of animals with only minor changes in LIN41 levels. First, *lin-41(bx37)* or *lin-41(bx42)* mutants probably only modestly overexpress LIN41, as they do not show any defects in development of the epidermis or the vulva (Del Rio-Albrechtsen et al., 2006). Second, male tail phenotypes of the temperature-sensitive *let-7(n2853)* mutant were assayed at 15 °C (Del Rio-Albrechtsen et al., 2006), its semi-permissive temperature, at which hermaphrodites do not burst through their vulva because LIN41 silencing by *let-7* is still partially functional (Ecsedi et al., 2015).

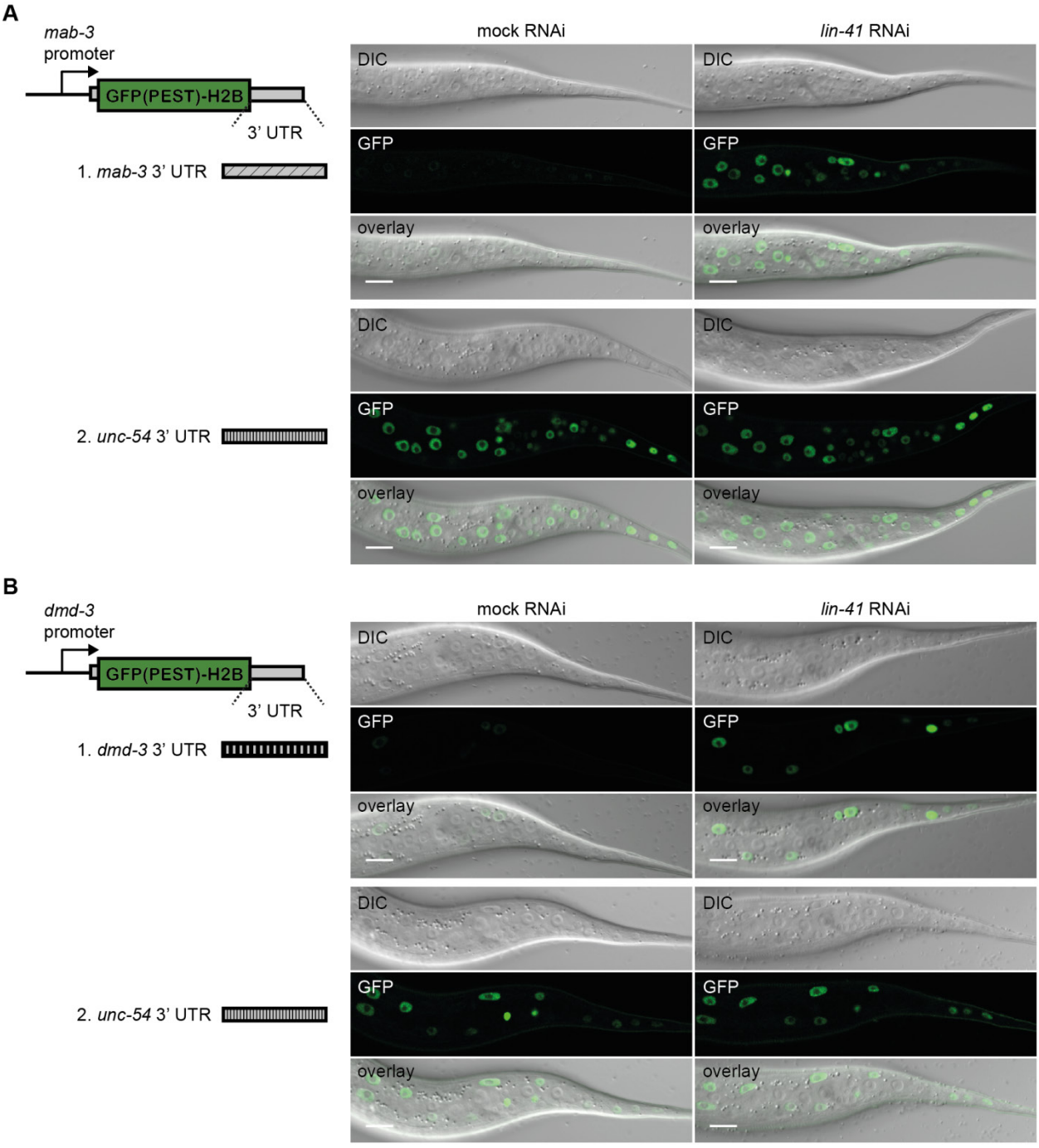


Figure 8. LIN41 regulates expression of *mab-3* and *dmd-3* in the male tail epidermis via their 3'UTRs
(A, B) Confocal images of the male tail epidermis in animals expressing nuclear-localized GFP(PEST)::H2B reporters, driven from with the *mab-3* (A) and *dmd-3* (B) promoters and fused to their endogenous 3'UTR sequences or the unregulated *unc-54* 3'UTR. Animals were grown for 20h at 25 °C on *lin-41* or mock RNAi bacteria. Scale bars: 10 µm.

Therefore, we observed the phenotype of male tails in *let-7(n2853)* mutants grown at 25 °C, the restrictive temperature. These males completely failed to retract their tails during the L4 stage, resulting in adult

worms with a pointed tail resembling hermaphrodite tails (Figure 9A), identical to those reported for *mab-3; dmd-3* double mutant males. The *let-7(n2853)* male tail phenotype was indeed caused by full dys-regulation of *lin-41*, as *lin-41(xe8)* mutant males phenocopied *let-7(n2853)* mutant males and as the *let-7(n2853)* phenotype was suppressed in the *lin-41(xe11)* mutant background (Figure 9A). From the interpretation of the observed phenotypes, we propose that the cell retraction defects in *lin-41(gf)* mutants can be explained by abnormal silencing of *mab-3* and *dmd-3* in L4 stage males.

2.5.4 LIN41 regulates female and male sexual organ development through a branched pathway

In summary, the four identified mRNA targets of LIN41 seem to act in two different pathways regulating the development of the female and male sexual organ, respectively (Figure 9B). On the one hand, the EGR transcription factor LIN-29A and its NAB-orthologous co-factor MAB-10 control proper morphogenesis of the vulva. On the other hand, the two DM-domain transcription factors DMD-3 and MAB-3 control cell retraction events to shape the male tail tip. LIN-29A and MAB-10, as well as DMD-3 and MAB-3, seem to have at least partially redundant roles in these processes, as single mutants have no obvious or only partially penetrant phenotypes. Since LIN41 silences expression of both redundant proteins in each pathway, elevated LIN41 levels in *lin-41(xe8)* mutants lead to fully penetrant phenotypes.

The two pathways seem to be completely separated. Vulva bursting seems to be independent of *mab-3* and *dmd-3*, because *mab-3; dmd-3* double mutants show no phenotype in hermaphrodites (Mason et al., 2008) and because depletion of *mab-3* and/or *dmd-3* by RNAi did not enhance the bursting phenotype of *lin-29(ΔA); mab-10(0)* mutant animals (data not shown). Moreover, the cell retractions to shape the male tail seem to be independent of LIN-29 activity (and thus likely MAB-10 activity), as the tail over-retraction phenotype of *lin-41(rf)* mutants is suppressed in *dmd-3* or *dmd-3;mab-3* double mutants, but not after depletion of LIN-29 by RNAi (Del Rio-Albrechtsen et al., 2006; Mason et al., 2008).

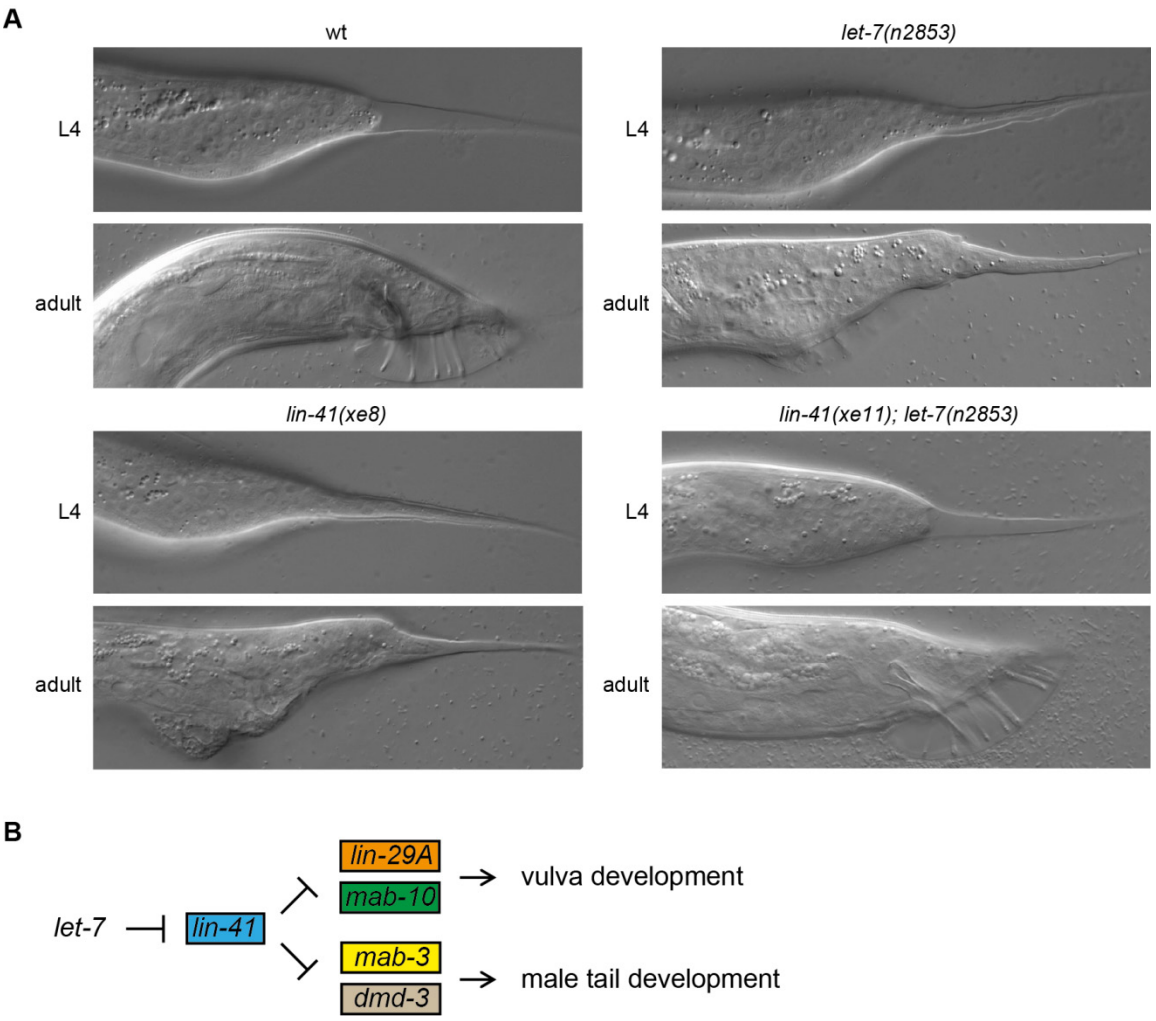


Figure 9. Sustained LIN41 expression in the L4 larval stage results in a complete failure in male tail retraction.
(A) Example micrographs of male tails in L4 larval stage and adult worms of the indicated genetic background. Worms were grown at 25 °C, the restrictive temperature for *let-7(n2853)* mutant animals. The male tail phenotypes were not quantified, but these images represent fully penetrant phenotypes observed in all assayed worms (n > 10).
(B) Model of the branched pathway downstream of the miRNA *let-7* regulating sexual organ morphogenesis in *C. elegans*.

2.6 The roles of LIN-29A, LIN-29B and MAB-10 in the L/A switch

The epidermal L/A switch is a term that summarizes four different events that occur during *C. elegans* development to transform the skin of worms from a larval state into an adult state (Figure 2B): seam cell exit from the cell cycle, terminal differentiation of the seam cells by cell fusion, alae secretion and termination of the molting cycle (Ambros, 1989). All four events might be coupled, for instance by a terminal differentiation program in seam cells that initiates cell fusion, with fused seam cells being incapable of going through cell divisions or initiating molts, but instead capable of producing alae. As another possibility, the exit from the cell cycle could signal the seam cells to fuse, produce alae and never initiate molts again. Alternatively, all L/A switch events could be independent from each other and could thus be regulated separately. An indication that the four L/A switch events could be coupled came from the discovery that *lin-29(0)* mutant animals do not go through any of the four L/A switch events (Ambros, 1984). Therefore, it seems possible that a change in general LIN-29 transcription factor activity as a sole input could regulate all four L/A switch events at once. However, we discovered that LIN41, itself a regulator of at least some L/A switch events (Slack et al., 2000), specifically controls expression of only one of the two LIN-29 isoforms, LIN-29A, and of its co-factor MAB-10. Therefore, we were interested whether the whole L/A switch with all four events is regulated by the *let-7*-LIN41 module and thus through LIN-29A and MAB-10. Alternatively, regulation of the LIN-29B isoform could also be necessary for controlling processes of the L/A switch. Hence, we studied the contributions of the *let-7*-LIN41 module, the two different LIN-29 isoforms and MAB-10 to the different L/A switch events.

2.6.1 LIN41 is the single key target of let-7 for control of seam cell self renewal

We first assayed the seam cell exit from the cell cycle, a process heavily studied for heterochronic mutants and reported to be misregulated in *lin-29*, *mab-10*, *let-7* and *lin-41* mutants (Ambros and Horvitz, 1984; Harris and Horvitz, 2011; Reinhart et al., 2000; Slack et al., 2000). As regulation of LIN41 alone accounts for the function of *let-7* in *C. elegans* vulval development (Ecsedi et al., 2015), we sought to explore whether LIN41 was also the single key target of *let-7* in controlling seam cell proliferation. To this end, we first examined seam cell proliferation in *lin-41(xe8)* and *let-7(n2853)* mutant animals. As also described in section 2.5, *lin-41(xe8)* animals lack a segment of the *lin-41* 3'UTR that contains the two functional *let-7* complementary sites (LCSs), thus specifically preventing *let-7*-mediated silencing of *lin-41* (Ecsedi et al., 2015). Consistent with dysregulation of *lin-41* being sufficient for seam cell overproliferation, we found that young adult *lin-41(xe8)* mutant animals displayed significantly elevated seam cell numbers relative to wild-types, comparable to the *let-7(n2853)* mutant phenotype (Figure 10A, B, see also Tables S5-S7 for

all seam cell quantifications). The increase in cell number results from failed termination of the self-renewal program, as *lin-41(xe8)* and *let-7(n2853)* mutant animals exhibit wild-type numbers of seam cells during the last larval stage, L4 (Figure 10B). Second, we analyzed seam cell numbers in *lin-41(xe11); let-7(n2853)* double mutant animals (Ecsedi et al., 2015). In these animals, the mutation in the *let-7* seed sequence leads to de-silencing of all *let-7* targets except for *lin-41*, whose two LCSs in the 3'UTR contain compensatory mutations to restore base pairing. Although *lin-41* is only partially down-regulated in this double mutant compared to wild-type animals (Aeschimann et al., 2017), even this incomplete restoration of repression sufficed to revert seam cell numbers to the lower, wild-type level (Figure 10A, B). Thus, LIN41 is the single key target of *let-7* for control of seam cell self-renewal.

2.6.2 Seam cell exit from the cell cycle is influenced by LIN-29A, LIN-29B and MAB-10

To investigate the roles of LIN-29A, LIN-29B and MAB-10 in different events of the L/A switch, we took advantage of mutant alleles we previously obtained by CRISPR-Cas9 (section 2.5, see Figure 6C for an illustration of the mutant alleles and Figure 6D for the specific allele names): A deletion of almost the whole *lin-29* locus and thus knocking out both isoforms, referred to as *lin-29(ΔAB)*, a deletion specifically disrupting expression of LIN-29A only, referred to as *lin-29(ΔA)*, and a deletion of almost the whole *mab-10* locus, referred to as *mab-10(0)*. As MAB-10 is a co-factor of both LIN-29 isoforms (Harris and Horvitz, 2011), the activity of both LIN-29A and LIN-29B may be modulated in *mab-10(0)* mutants. MAB-10 itself is thought not to directly bind to DNA, suggesting that *lin-29(ΔAB)* mutants lack any LIN-29- or MAB-10-related activity. Consequently, in the *lin-29(ΔA)* genetic background, LIN-29B can still act together with its co-factor MAB-10, while the *lin-29(ΔA);mab-10(0)* mutants are left with LIN-29B, acting without its co-factor MAB-10. With this set of mutants, we could not study worms left with only LIN-29A activity (with or without MAB-10). This is due to our lack of knowledge about elements (e.g. within the promoter) specific to *lin-29B* that could be mutated without affecting expression of *lin-29A,* currently precluding us from obtaining a specific knockout of *lin-29B*. To study isoform-specific differences in events of the L/A switch, we thus investigated the differences between *lin-29(ΔA)* and *lin-29(ΔAB)*, with or without the additional *mab-10(0)* mutation.

Since sustained LIN41 expression leads to a failure in seam cell cycle exit, we examined the functions of the LIN41 targets LIN-29A and MAB-10 in this process. Seam cell numbers in young adults of either *lin-29A(0)* or *mab-10(0)* single mutants were unchanged from the wild-type situation (Figure 10C). By contrast, *mab-10(0) lin-29A(0)* double mutant animals displayed seam cell overproliferation (Figure 10C) comparable to that seen in *let-7(n2853)* or *lin-41(xe8)* mutant animals (Figure 10B). We conclude that

these two proteins are a major, perhaps sole, output of the *let-7*–LIN41 regulatory module for control of seam cell self-renewal.

However, in *let-7(n2853)*, *lin-41(xe8)* and *mab-10(0) lin-29A(0)* mutants, only a bit more than half of the animals harbored at least 25 seam cells instead of the usual 16 at the young adult stage, right before many died by bursting (Figure 10B,C). Therefore, this seems to be a partially penetrant phenotype, also because young adult *lin-29(ΔAB)* and *lin-29(ΔAB); mab-10(0)* mutant worms without exception had 25 or more seam cells (Figure 10C). This suggests that the incomplete penetrance of this phenotype in *let-7(n2853)*, *lin-41(xe8)* and *lin-29(ΔA); mab-10(0)* mutants was due to the presence of LIN-29B.

As described above, we did not observe additional seam cells in young adult *mab-10(0)* animals (Figure 10C). By contrast, in a previous study, analysis of two isolated *mab-10* point mutants, *mab-10(n5117)* and *mab-10(n5118)* (see also Figure 6C), indicated that MAB-10 is required for exit from the cell cycle in seam cells (Harris and Horvitz, 2011). However, overproliferation of seam cells was only observed in older adults, earliest at about 14 hours after the molt (Harris and Horvitz, 2011). Therefore, we additionally analyzed the number of seam cells in older adults of each genotype, at the stage when they contain the first embryo. Indeed, we confirmed extra seam cell divisions in older *mab-10(0)* mutant adults (Figure 10C). At this developmental stage, the absence of LIN-29A alone did not yield extra seam cells, while the absence of both LIN-29A and MAB-10 led to a completely penetrant overproliferation phenotype, as observed for *lin-29(ΔAB)* mutant worms at the young adult stage (Figure 10C). Seam cells of *lin-29(ΔAB)* mutants, with or without additional mutation of *mab-10*, were less numerous in older than younger adults, with some animals even harboring the normal 16 seam cells (Figure 10C). This suggested that their extra seam cell divisions were asymmetric, as for those during late larval development, and that one of the daughter cells from each division had joined the hyp7 syncytium, thus losing the expression of the seam cell specific marker. Taken together, these results indicated that complete absence of LIN-29 activity results in rapid extra seam cell divisions right after the L4-to-adult molt, with additional seam cell nuclei joining hyp7 as they grow to older adults. LIN-29B activity alone, in absence of LIN-29A and MAB-10, is not sufficient to stop seam cell proliferation, but can delay or slow down the extra cell division, resulting in the observed fully penetrant overproliferation phenotype in older adults. However, together with MAB-10, LIN-29B is sufficient to terminate self-renewal activity in seam cells, suggesting MAB-10 activity is crucial for this event. Consistently, LIN-29A and LIN-29B without their co-factor MAB-10 are also not able to completely terminate seam cell divisions, as some seam cells keep dividing, although even later in adulthood compared to *lin-29(ΔA); mab-10(0)* mutant animals. The vulval bursting of *let-7(n2853)* and *lin-41(xe8)* mutants at the young adult stage precluded us from phenotyping them at the older adult stage.

Nonetheless, considering that in these mutants, LIN-29A and MAB-10 expression are silenced, we suspect we would observe a fully penetrant overproliferation phenotype later in adulthood if they had survived.

2.6.3 Expression of LIN-29B is sufficient for seam cell fusions

In order to study whether the seam cell proliferation phenotypes correlate with seam cell fusion defects, we analysed the fusion of seam cells in the different mutant backgrounds using *ajm-1::mCherry*, a marker of adherens junctions that surround all epidermal cells. It was previously shown that seam cells failed to fuse in adult *lin-29(n546)* mutant animals (Bettinger et al., 1996). The same was observed in a strain overexpressing LIN41 (Slack et al., 2000). By contrast, seam cell fusion was reported to normally occur in *let-7(mn112)* null mutant animals (Hunter et al., 2013). Taken together, this would suggest that LIN41 and LIN-29 are involved in seam cell fusion, independently of regulation by *let-7*. This seemed rather unlikely, as LIN41 is highly overexpressed in the epidermis of *let-7* mutants. Thus, it has yet been unclear if and how seam cell fusion is controlled by the *let-7*-LIN41 pathway and, consequently, what the roles of LIN-29A, LIN-29B and MAB-10 in this process could be. We observed that seam cell fusion in wild-type worms occured at the mid L4 stage, after the last regular seam cell division and before the last molt. Therefore, we assayed the *ajm-1::mCherry* marker in late L4 stage worms, when seam cell numbers were still normal in all mutant backgrounds, and in young and older adults, when seam cell overproliferation occurred in certain genetic backgrounds. At the late L4 stage, wild-type, *mab-10(0)* and *lin-29(ΔA)* animals had normally fused seam cells (n=20 each, Figure 10D). By contrast, the seam cells of late L4 stage *lin-29(ΔAB)* and *lin-29(ΔAB); mab-10(0)* mutant worms had failed to fuse in all analyzed animals (n=20 each, Figure 10D and data not shown), confirming that some LIN-29 activity is necessary for seam cell fusion (Ambros and Horvitz, 1984; Bettinger et al., 1996). However, suprisingly, seam cells had normally fused in all observed late L4 stage *lin-29(ΔA); mab-10(0)* mutants (n=20, Figure 10D), as well as *let-7(n2853)* and *lin-41(xe8)* mutant animals (n=20 each). We conclude that LIN-29B alone, without its co-factor MAB-10, is sufficient for seam cell fusion, and thus seems to be the main driver of this event of the L/A switch. Furthermore, this result confirmed the data obtained with the *let-7(mn112)* mutant (Hunter et al., 2013), but was in contradiction to the data with the strain overexpressing LIN41 (Slack et al., 2000).

2.6.4 Extra nuclear divisions divide up the seam syncytium

If there is one syncytium containing all seam cell nuclei in late L4 stage *let-7(n2853)*, *lin-41(xe8)* and *lin-29(ΔA); mab-10(0)* animals, what we described as extra seam cell divisions in young adults could in reality

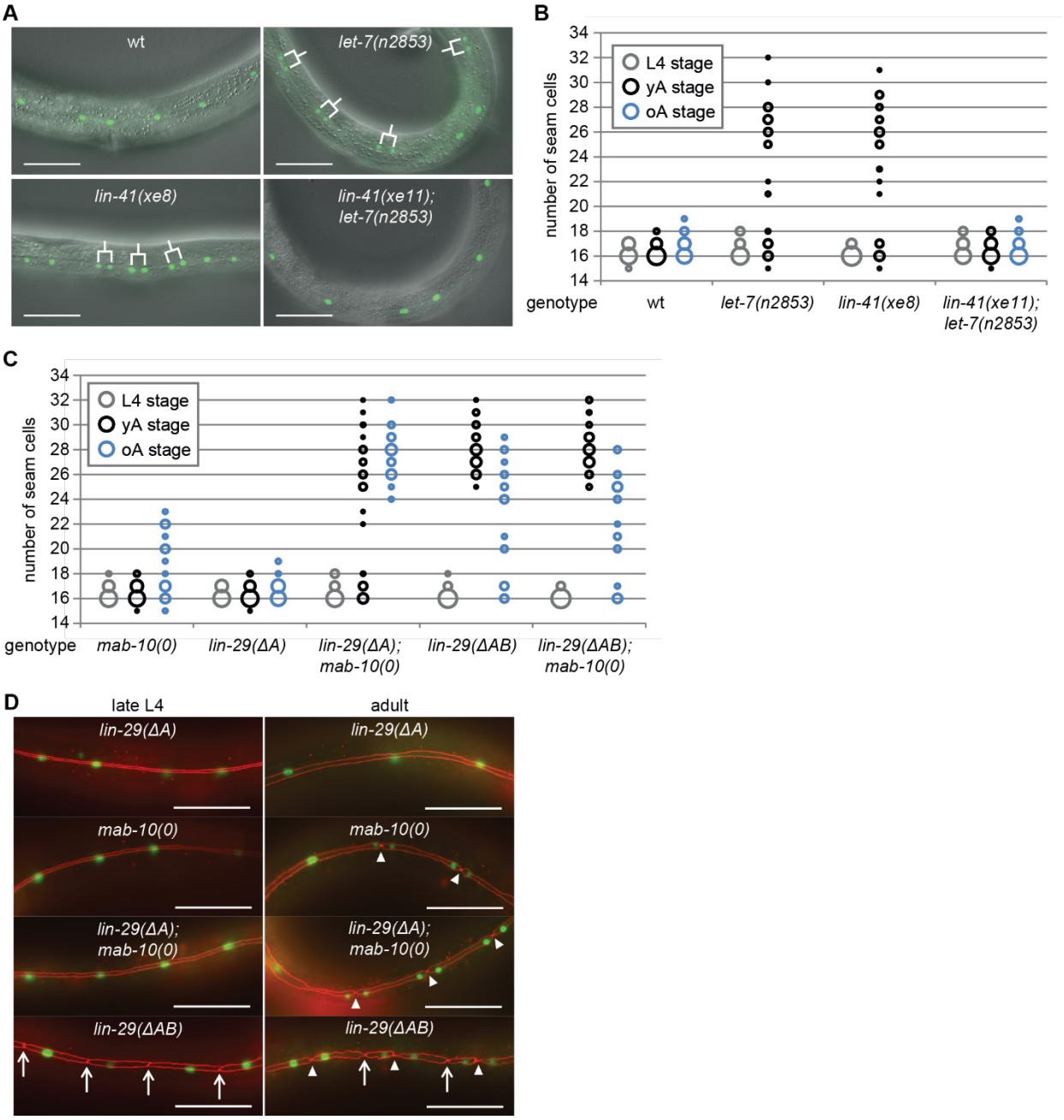


Figure 10. The roles of *let-7*, LIN41, LIN-29A, LIN-29B and MAB-10 in proliferation and fusion of seam cells
(A) Example micrographs of young adult worms expressing transgenic *scm::gfp* to visualize seam cells (Koh and Rothman, 2001). Branched lines indicate seam cells originating from extra cell divisions. Scale bars: 50 µm.
(B, C) L4 larval stage, young adult (yA) and older adult (oA) worm seam cell number quantifications in indicated genetic backgrounds. Areas of bubbles represent the percentage of worms with the respective number of seam cells. n=20 for L4, n>50 for yA, n=20 for oA worms.
(D) Example micrographs of animals expressing transgenic *scm::gfp* to visualize seam cells (out of focus) and *ajm::mCherry* to visualize epidermal cell junctions (in focus). Imaged is the epidermis of late L4 worms for all genotypes and of young adult worms for all genotypes but *mab-10(0)*, where the epidermis of an older adult is shown to illustrate the extra seam cell division. Arrows indicate cell boundaries due to a failure of seam cell fusion during the L4 stage, arrowheads indicated newly established cell boundaries in adult worms at the sites of extra nuclear divisions. Scale bars: 50 µm.

be endoreduplications with extra nuclear divisions within the syncytium. However, the next surprise came when we observed the *ajm-1::mCherry* marker in young adult worms. The seam cell nuclei in *let-7(n2853)*, *lin-41(xe8)* and *lin-29(ΔA); mab-10(0)* mutants that underwent an extra nuclear division again established cell junctions between the newly borne nuclei, dividing up the syncytium (Figure 10D, and data not shown for *let-7(n2853)* and *lin-41(xe8)*). These interruptions of the syncytium specifically occurred at the locations of nuclear divisions, something that was also clearly apparent in older adults of *mab-10(0)* mutants, where only few seam cell nuclei divided (Figure 10D). In *lin-29(ΔAB)* and *lin-29(ΔAB); mab-10(0)* mutants, new cell junctions were also established after nuclear division as in a normal seam cell division of earlier larval stages, whereas in wild-type and *lin-29(ΔA)* worms, all seam nuclei continued to be fused (Figure 10D). In summary, seam cell fusion can occur before and independently of the later presence of extra seam cell nuclei, suggesting that exit from the cell cycle and the ability to fuse are independent events of the L/A switch. By contrast, extra nuclear divisions seem to re-establish cell-to-cell junctions in between the newly borne nuclei. Notably, the latter explains the published observation that overexpression of LIN41 results in a failure of seam cell fusion. As these worms were assayed as young adult animals, the re-established cell junctions were mis-interpreted as a consequence of failed seam cell fusions in the L4 stage (Slack et al., 2000).

2.6.5 Wild-type alae formation is dependent on expression of both LIN-29 isoforms

The commonly used readout for terminal differentiation of seam cells is the presence of cuticular alae. These ridges are formed on the adult, but not the L4 larval stage cuticle, and are therefore thought to be secreted by terminally differentiated and fused seam cells. Our results would suggest that the connection between cell fusion and alae secretion is less clear, as alae have been reported to be absent in *let-7(n2853)* and LIN41-overexpressing animals (Reinhart et al., 2000; Slack et al., 2000), where seam cell fusion normally occurs. We analysed alae formation in young and older adult animals of the different genetic backgrounds (Figure 11, see also Tables S8-S9 for all alae quantifications). As expected from the literature (Harris and Horvitz, 2011), alae formation was unaffected in *mab-10(0)* mutants at any stage in adulthood, resulting in clearly visible (“strong”) cuticular ridges along the whole worm body, as in wild-type animals (Figure 11). By contrast, no alae were found in young or older adult *lin-29(ΔAB)* and *lin-29(ΔAB); mab-10(0)* animals, confirming that LIN-29 activity is necessary for alae secretion (Figure 11) (Ambros and Horvitz, 1984). The absence of LIN-29A alone yielded weaker alae structures covering at least 50 % of the body length of young adult and the whole body of older adult animals. In contrast to seam cell fusion and exit from the cell cycle, depletion of LIN-29A alone thus did have an effect on alae secretion. Interestingly, and in striking contrast to seam cell fusion and exit from the cell cycle, co-depletion of MAB-10 did not

enhance the *lin-29(ΔA)* phenotype. The wild-type alae of *mab-10(0)* animals and the identical alae phenotypes of *lin-29(ΔA); mab-10(0)* and *lin-29(ΔA)* mutant animals suggest that MAB-10 does not have a function in alae secretion. In contrast to MAB-10 and similar to its role in seam cell fusion, LIN-29B seems to be sufficient for formation of alae. Nevertheless, in order to produce strong, wild-type alae structures, both LIN-29 isoforms need to be present. Consistent with these interpretations, we also observed weak alae structures in some young adult *let-7(n2853)* and *lin-41(xe8)* animals and speculate that we would find weak, but complete alae if they survived into later adulthood. However, as we could not find alae structures on more than 50 % of young adult *let-7(n2853)* and *lin-41(xe8)* animals, it remains possible that in addition to LIN-29A, other LIN41 targets may be involved in alae formation.

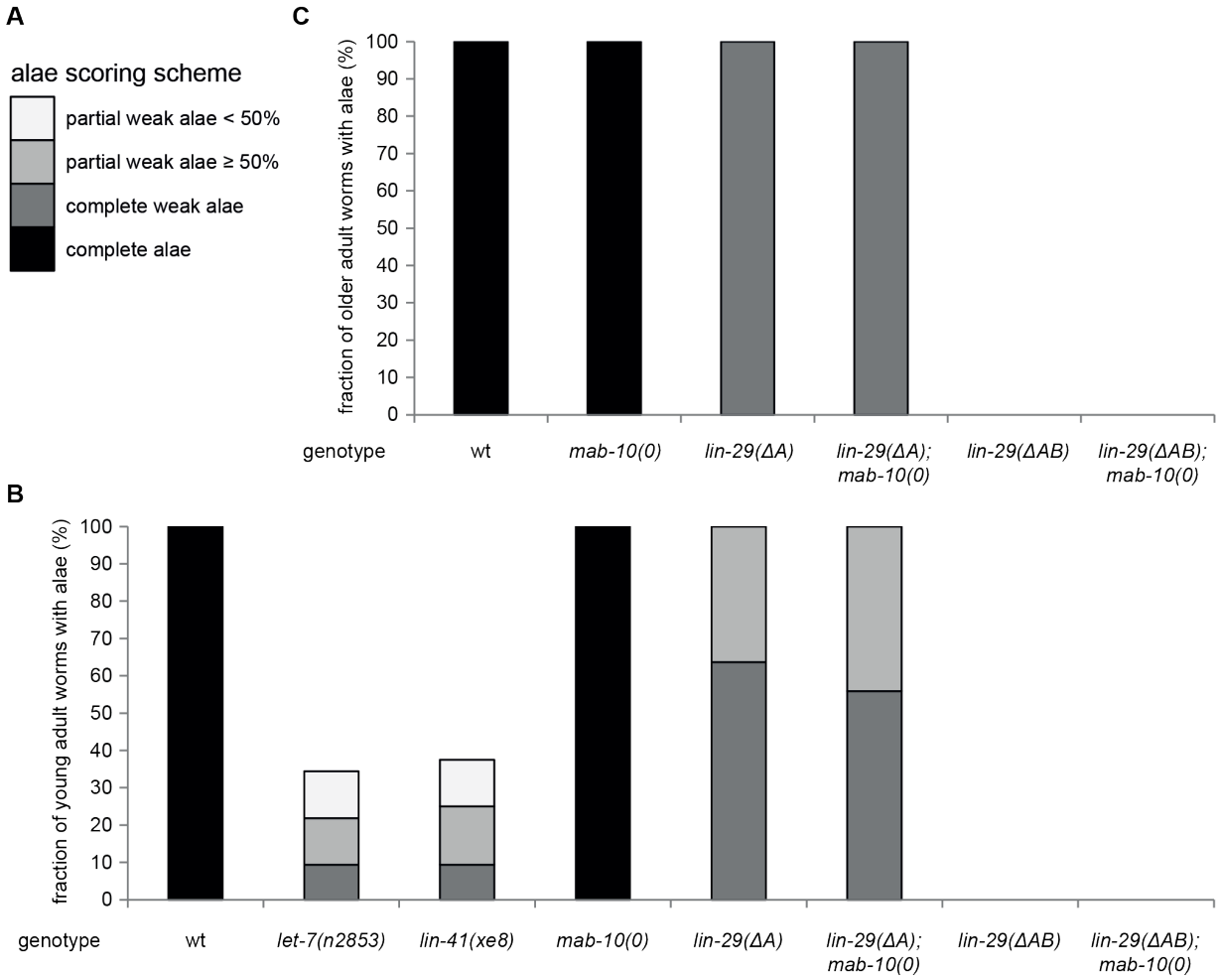


Figure 11. Alae formation depends on LIN-29A and LIN-29B, but not on MAB-10
(A) The scoring scheme for alae structures used for (B, C).
(B, C) Quantification of different alae structures on the cuticules of young adult (B) and older adult (C) worms of indicated genotypes. n≥32 for young adults, n=20 for older adults.

2.6.6 Different L/A switch events require different LIN-29 activities

In summary, although the complete absence of any LIN-29 activity results in failure of all the L/A switch events, our analysis of the L/A switch events in the different genetic backgrounds reveals that seam cell exit from the cell cycle, seam cell fusion and alae secretion have different requirements for the activities of LIN-29A, LIN-29B and MAB-10 (Figure 12). While LIN-29B alone is sufficient for seam cell fusion, LIN-29B and MAB-10 together are sufficient to stop seam cell proliferation and the two LIN-29 isoforms together in absence of MAB-10 are sufficient for wild-type alae formation. As the different events of the L/A switch do not take place simultaneously, it will be interesting to find if the worm regulates the timing of these events through isoform-specific regulation of LIN-29. Thereby, *let-7* and LIN41 may regulate the timing of seam cell exit from the cell cycle through LIN-29A, but seam cell fusion might be regulated through a different pathway eventually resulting in control of LIN-29B expression.

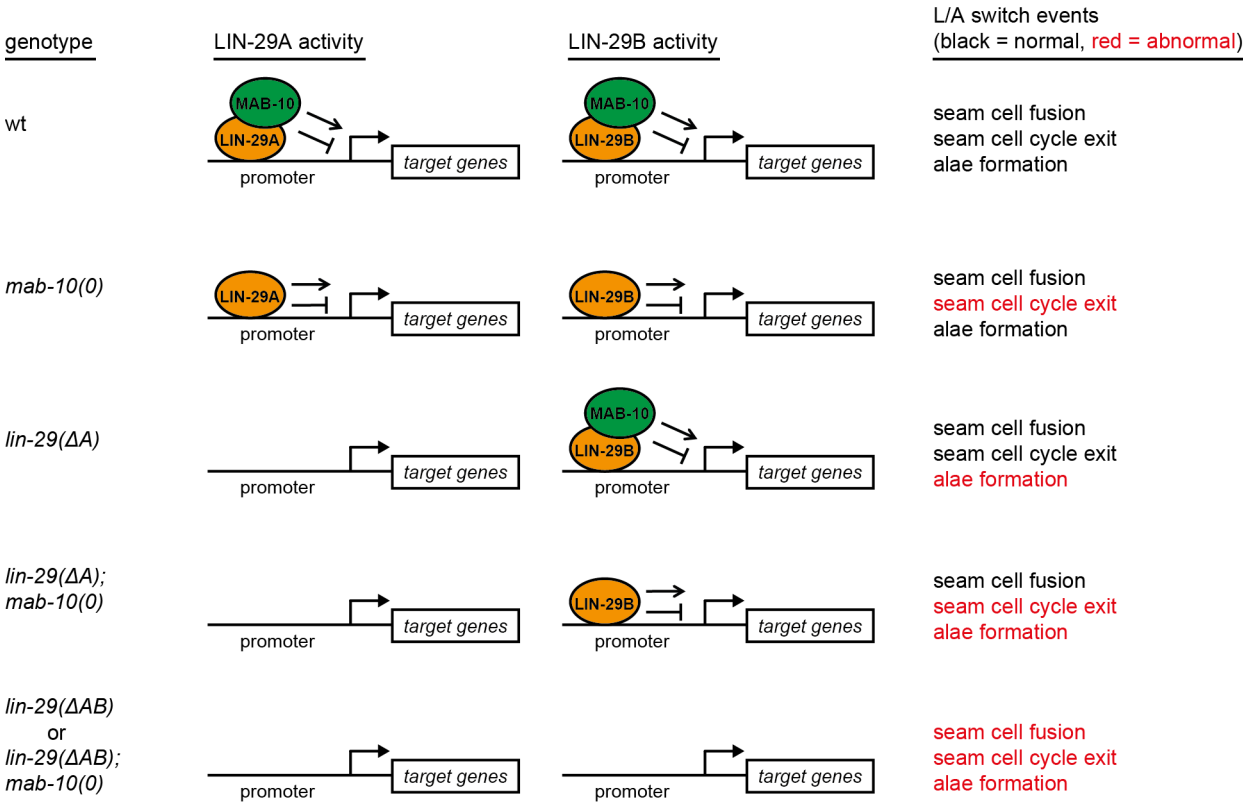


Figure 12. Summary of the roles of LIN-29A, LIN-29B and MAB-10 in events of the L/A switch.
Seam cell fusion, seam cell exit from the cell cycle and alae secretion are differently affected in the different genetic backgrounds, probably due to different requirements of LIN-29 isoform-specific activities. LIN-29A and LIN-29B, together with MAB-10, may enhance or suppress transcriptional activity on their target genes, as both activities have been described for their mammalian counterparts (Sevetson et al., 2000; Svaren et al., 1998). It is currently unknown how LIN-29A and LIN-29B differ in regulating their specific target genes, and how MAB-10 can modulate the activities of the two LIN-29 isoforms.

2.7 The two LIN-29 isoforms are not redundant

When *lin-29* was first mapped, sequenced and detected at the mRNA level by Rougvie and Ambros in 1995, it became clear that there were two major mRNA isoforms, to which they referred to as *lin-29A* and *lin-29B* (Rougvie and Ambros, 1995). Soon after, Bettinger *et al.* reported different expression patterns for the two isoforms, based on promoter fusion assays (Bettinger et al., 1996). However, in a follow-up study, the same authors suggested that the two isoforms are expressed in the same cells and at the same time (Bettinger et al., 1997). This conclusion was based on similar LIN-29 protein expression patterns in wild-type worms and worms with the *lin-29(ga94)* mutation, thought to specifically interfere with LIN-29A but not LIN-29B expression due to a premature stop codon in exon 5. Thus, the *lin-29A* mRNA was thought to be degraded through NMD, while *lin-29*B was proposed to be translated from an open reading frame starting immediately downstream of the introduced stop codon. As in *lin-29(ga94)* mutant animals, immunostainings revealed lower than wild-type LIN-29 accumulation without noticeable alteration of the expression pattern, LIN-29A and LIN-29B were proposed to redundantly contribute to the same functions (Bettinger et al., 1997). This idea seemed valid, as the presumed loss of LIN-29A in *lin-29(ga94)* mutants caused vulval protrusions (Pvl), a phenotype that could be rescued by overexpression of LIN-29B in *lin-29(0)* mutants (Bettinger et al., 1997). Taken together, and because of the uncertainty that the cloned promoters contained all elements influencing expression of the LIN-29 isoforms (Bettinger et al., 1996), LIN-29A and LIN-29B were since considered to be redundant and thus have not been individually studied again.

We observed clear differences in epidermal and vulval phenotypes between *lin-29(0)* and *lin-29A* mutants (sections 2.5 and 2.6), which could hint towards different roles of the two isoforms, but could also reflect different thresholds of general LIN-29 activity necessary for the different functions in development. However, when investigating the function of LIN41 in *C. elegans* somatic development, we found that the LIN-29A isoform is specifically regulated by LIN41 through an elaborate mechanism (Aeschimann et al., 2017). We therefore hypothesized that this regulation, instead of merely serving to partially reduce LIN-29 activity, could be a mechanism to distinctly regulate LIN-29 activity in different tissues or at different time points in development. Hence, although in contradiction to the results obtained with the *lin-29(ga94)* mutant (Bettinger et al., 1997), we wondered if LIN-29A and LIN-29B are expressed at different developmental times and/or in different tissues.

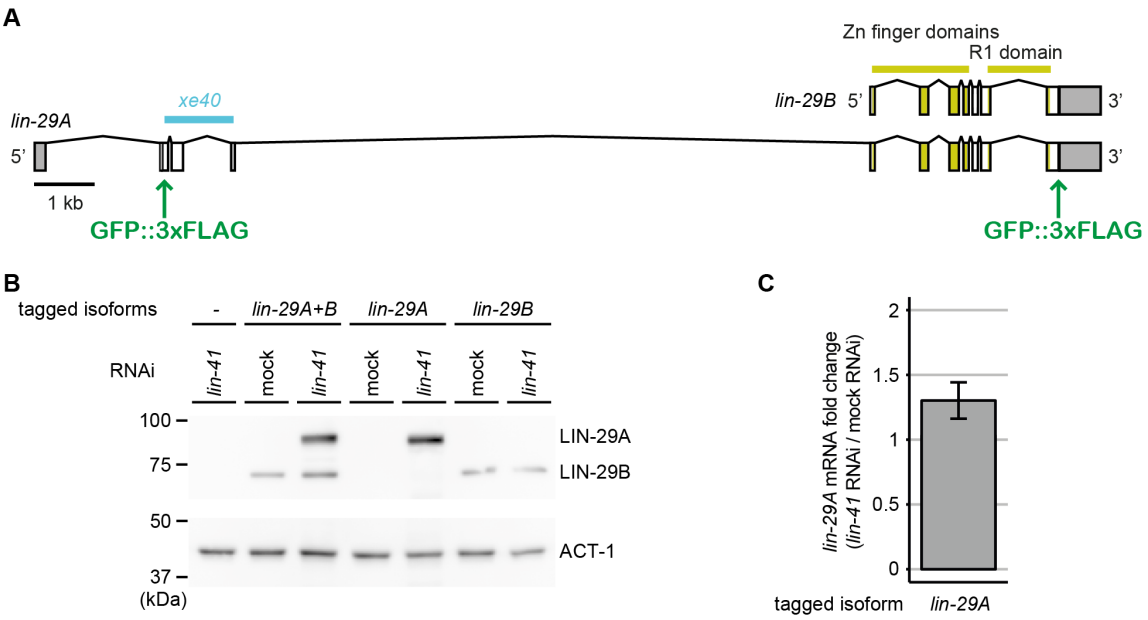


Figure 13. Isoform-specific tagging of LIN-29A and LIN-29B using CRISPR-Cas9
(A) Illustration of the locations where a GFP-3xFLAG tag was inserted into the *lin-29* genomic locus. The GFP-3xFLAG tag at the N-terminus of *lin-29A* specifically tags LIN-29A, while the GFP-3xFLAG tag at the shared C-terminus of both isoforms specifically tags LIN-29B if combined with a *lin-29(xe40)* (*lin-29(ΔA)*) genetic background.
(B) Western blot analysis to detect endogenous GFP::3xFLAG-tagged LIN-29A and LIN-29B proteins using an anti-FLAG antibody. ACT-1 was detected as a loading control. Animals were grown for 20h at 25 °C on *lin-41* or mock RNAi bacteria. Worm strains: *lin-29(xe61*[*lin-29::gfp::3xflag*]*)*, depicted as "*lin-29A+B*", *lin-29(xe63*[*gfp::3xflag::lin-29A*]*)*, depicted as "*lin-29A*", *lin-29(xe65*[*lin-29A(xe40)::gfp::3xflag*]*)*, depicted as "*lin-29B*".
(C) RT-qPCR analysis to verify that the change in LIN-29A protein expression detected on the Western blot (lanes with specifically tagged *lin-29A* in (B)) is not accompanied by a change in *lin-29A* mRNA levels. Depicted is the fold change of *lin-29A* mRNA levels (normalized by *act-1* mRNA levels). Worms were grown as described above (B). n = 3 biological replicates, data as mean ± s.e.m. Worm strain: *lin-29(xe63*[*gfp::3xflag::lin-29A*]*)*, depicted as "*lin-29A*".

2.7.1 Isoform-specific tagging of LIN-29A and LIN-29B using CRISPR-Cas9

In order to study the expression patterns of LIN-29A and LIN-29B, we separately tagged each LIN-29 isoform with a GFP::3xFLAG tag at its endogenous locus using CRISPR-Cas9-mediated genome editing (Dickinson et al., 2015). On the one hand, we placed a GFP::3xFLAG tag at the N-terminus of LIN-29A, resulting in specific tagging of LIN-29A (Figure 13A). On the other hand, we placed a GFP::3xFLAG tag on the C-terminus of the LIN-29 ORF, in the *lin-29(ΔA)* mutant background (Figure 13A). Thus, although the C-terminus is shared between the two isoforms, only GFP::3xFLAG-tagged LIN-29B is expressed due to the *lin-29A*-specific mutation. To verify that tagging was specific to the isoforms, we performed Western blotting of L3-stage worms containing tagged LIN-29A or LIN-29B, in the presence and absence of LIN41 (Figure 13B). For comparison, we performed the same analysis on unmodified wild-type worms and worms with a 3xFLAG-GFP tag on both isoforms. Because LIN41 represses LIN-29A (Aeschimann et al., 2017), only LIN-29B can normally be detected in L3-stage worms. By contrast, upon depletion of LIN41, LIN-29A accumulates to even higher levels than LIN-29B ((Aeschimann et al., 2017) and Figure 13B).

Whereas only the accumulation of LIN-29A could be detected in worms with the GFP::3xFLAG tag at the N-terminus of LIN-29A, only the unchanged LIN-29B expression was observed in *lin-29(ΔA)* worms with a GFP::3xFLAG tag at the C-terminus, confirming that we managed to separately tag each LIN-29 isoform (Figure 13B). As expected, the massive accumulation of N-terminally tagged LIN-29A protein upon LIN41 depletion was due to regulation on the translational level, as *lin-29A* RNA levels did not change to a comparable extent (Figure 13C). Notably, the GFP::3xFLAG tagged LIN-29A protein seemed to be fully functional, as the animals expressing GFP::3xFLAG::LIN-29A did not show any of the phenotypes associated with *lin-29A* mutants, whereas the phenotype of *lin-29(ΔA)* animals expressing LIN-29B::GFP::3xFLAG resembled that of *lin-29(ΔA)* mutants without inserted tag, but with a slightly increased occurrence of Pvl and burst animals, suggesting a partial decrease in functionality of the protein.

2.7.2 Epidermal expression patterns of LIN-29A and LIN-29B are different

To study the specific expression of LIN-29A and LIN-29B over development in different tissues, we used confocal imaging to detect nuclear accumulation of GFP in the worm lines with individually tagged LIN-29 isoforms, as well as the worm line with both isoforms tagged. Due to the crucial functions of LIN-29 in the L/A switch (section 2.6), we first observed the protein isoform expression over development in the epidermis. We found the earliest epidermal nuclear GFP signal in seam cells of late L3 stage worms. Strikingly, this earliest seam cell accumulation was clearly specific to LIN-29B, and LIN-29A was not expressed in the epidermis at this time (Figure 14A). Only in mid L4 stage animals, at the time LIN41 relieves LIN-29A from repression (Aeschimann et al., 2017), LIN-29A accumulated in the epidermis, in seam cells and in the epidermal syncytium hyp7 (Figure 14B). At that stage, weak LIN-29B was also detected in hyp7, and as it was strongly expressed in the seam cells migrating away from the midline just after their final division (Figure 14B), it seems possible that after fusion to hyp7, LIN-29B is distributed to all hyp7 nuclei by shuttling between nuclei and cytoplasm. Consistent with this, further LIN-29B accumulation in hyp7 was observed in the late L4 stage (Figure 14C). Both LIN-29A and LIN-29B expression in seam and hyp7 cells persisted into adulthood (data not shown). In summary, these results demonstrate different epidermal expression patterns for LIN-29A and LIN-29B and thus contradict the hypothesis that these isoforms are redundant (Bettinger et al., 1997). Of note, and in contrast to the previous report (Bettinger et al., 1997), LIN-29B accumulated in seam cells of worms just prior to the L3-to-L4 molt and thus before the final division of seam cells. Consequently, LIN-29B accumulation alone seems not to be sufficient for seam cell exit from the cell cycle, consistent with the data of section 2.6. In addition, LIN-29A accumulates only in the mid L4 stage, when fused seam cells are already observed, supporting our previous conclusion that LIN-29B activity is sufficient for seam cell fusion.

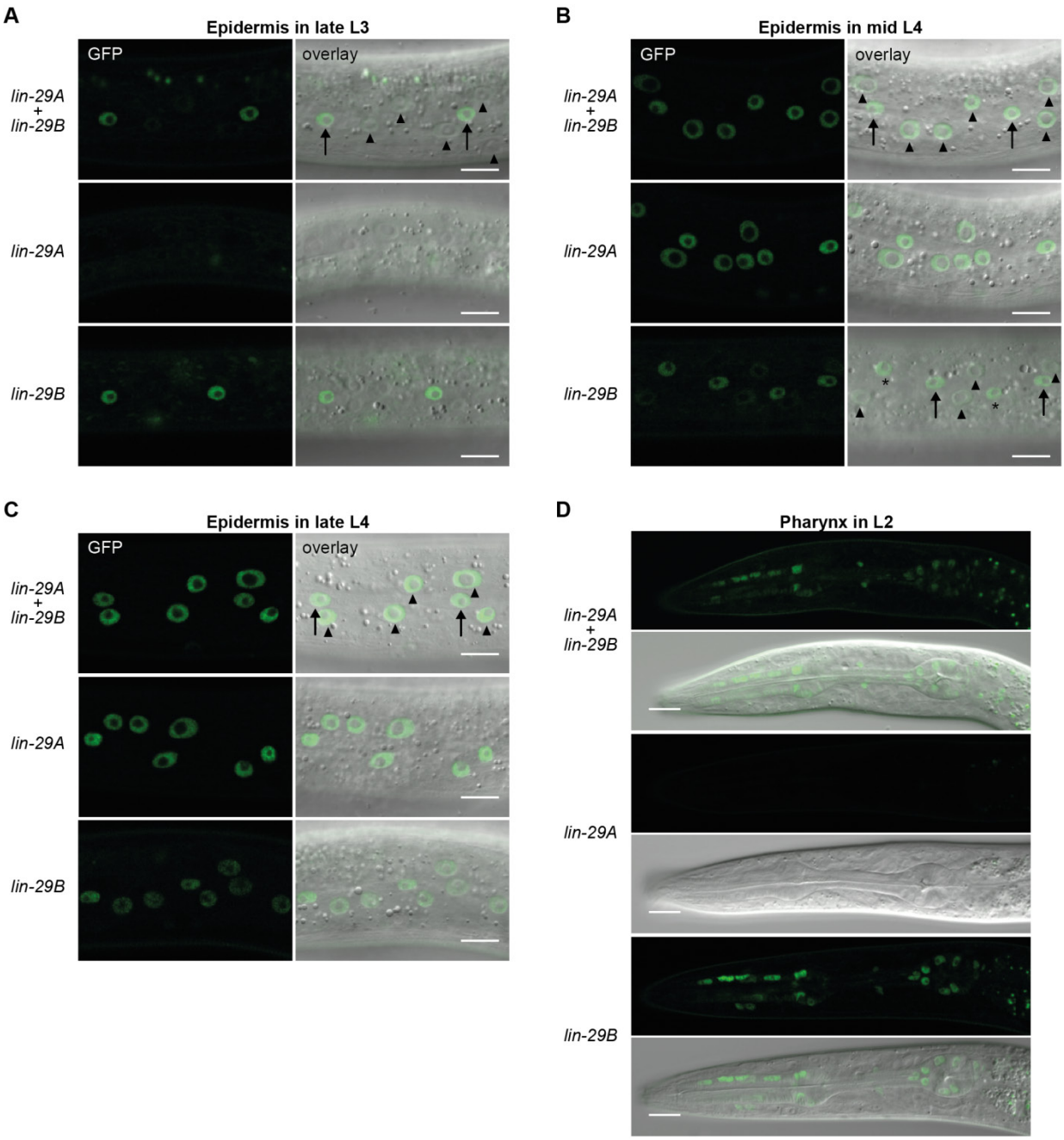


Figure 14. Expression patterns of LIN-29A and LIN-29B in the epidermis and the pharynx
(A-C) Confocal images of endogenously tagged LIN-29 protein isoforms in the epidermis of animals at the indicated developmental stages. Scale bars: 10 µm. At the late L3 stage (A), LIN-29B accumulates specifically in the nuclei of seam cells (arrows), and not of hyp7 (arrowheads). In mid (B) and late (C) L4 stage animals, both LIN-29A and LIN-29B are expressed in seam cell (arrows) and hyp7 (arrowheads) nuclei. LIN-29A is more strongly expressed in hyp7 nuclei compared to LIN-29B, while the expression level of both isoforms in seam cell nuclei is more similar. At the mid L4 stage (B), after the last seam cell division, LIN-29B is clearly detected in both the posterior daughter cells (arrows) and the anterior daughter cells (asterisks) that fuse with hyp7.
(D) In the pharynx, LIN-29B is expressed throughout larval and adult development (shown are L2 stage worms). Scale bars: 10 µm.

2.7.3 Expression patterns of LIN-29A and LIN-29B in other tissues

LIN-29A and LIN-29B were also clearly expressed in a non-redundant manner in tissues other than the epidermis. We observed the most striking differences in the head, tail and vulval regions. In the head region, we detected specific expression of LIN-29B in the metacorpus and the terminal bulb of the pharynx (Figure 14D). Weak expression of LIN-29B in these tissues was already detected in L1 worms, increased over development and persisted into adulthood.

Expression patterns in the region of the future vulva were particularly interesting during the L3 stage. At that time, the sex myoblasts (SMs) and vulval precursor cells (VPCs) go through several rounds of cell divisions, in order to prepare the morphogenesis of the vulva and uterus that mainly occurs in the L4 stage. During these processes, vulval and uterine cell fates are largely specified by signalling from a single cell called the anchor cell (AC). Approximately at the L2-to-L3 molt, the anchor cell started to strongly and specifically express LIN-29A (Figure 15A,B,D). At the same time, LIN-29B was weakly expressed in the two SMs (Figure 15A), an expression that became much stronger in the mid and late L3 stage, especially in the daughters of the SMs after their first cell division (Figure 15C). Similarly, the three VPCs P5.p-P7.p showed specific expression of LIN-29B that was weak in the three mid L3 stage VPCs but much stronger in their 6 daughters, shortly after they were born (Figure 15B). After a further VPC cell division, in worms of the later L3 stage, LIN-29B expression was accompanied by accumulation of LIN-29A, specifically in the granddaughters of P5.p and P7.p, but not in those of P6.p (Figure 15D).

As in the head region in early larval stages, we detected expression of LIN-29B, but not LIN-29A, in the tail region of L3 stage worms. During the mid L3 stage, LIN-29B accumulated in two rectal cells (B and hyp12/P12.pa) (Figure 16A), while in late L3 worms, four additional cells (K.a, K', F and U) in the rectal region also started to accumulate LIN-29B, from which K.a and F are shown in Figure 16B (K' and U are in a different focal plane).

2.7.4 LIN-29A and LIN-29B are regulated at the level of translation and mRNA abundance, respectively

While the obtained tagged lines provide the first tool to study the distinct LIN-29 isoforms at the protein level, they had been studied in whole-worm lysates at the mRNA level at the time of their discovery (Rougvie and Ambros, 1995). Northern blot analysis with RNA from worms at different developmental stages showed a largely similar developmental expression pattern for the two isoforms at the mRNA level (Figure 17A, (Rougvie and Ambros, 1995)). Both isoforms were weakly expressed from the L1 stage on and greatly increased in abundance over larval development, peaking at the L4 stage. In young and older adults, the mRNA levels were again decreased, to similar levels as in the L1 stage. Except for at the L4

stage, the *lin-29A* mRNA seemed to be higher in abundance compared to the *lin-29B* mRNA. In order to perform an according experiment to detect the LIN-29 isoforms on the protein level, we performed Western blotting of worms with a GFP::3xFLAG tag on both isoforms at different developmental stages. In striking contrast to the Northern blot analysis by Rougvie *et al*., we did not detect any LIN-29A protein in L1 or L2 stage worms, before detecting a weak accumulation in L3 stage and a strong accumulation in L4 stage worms. In combination with our previous results, we conclude that LIN41 inhibits the translation of LIN-29A during the first three larval stages. Contrary to LIN-29A, the expression of the LIN-29B protein over development largely correlated with its mRNA abundance. LIN-29B expression over development thus seems to be mainly regulated at the mRNA level. Notably, we could detect both LIN-29 protein isoforms in young and old adults despite low mRNA levels, suggesting that the LIN-29 proteins are rather stable.

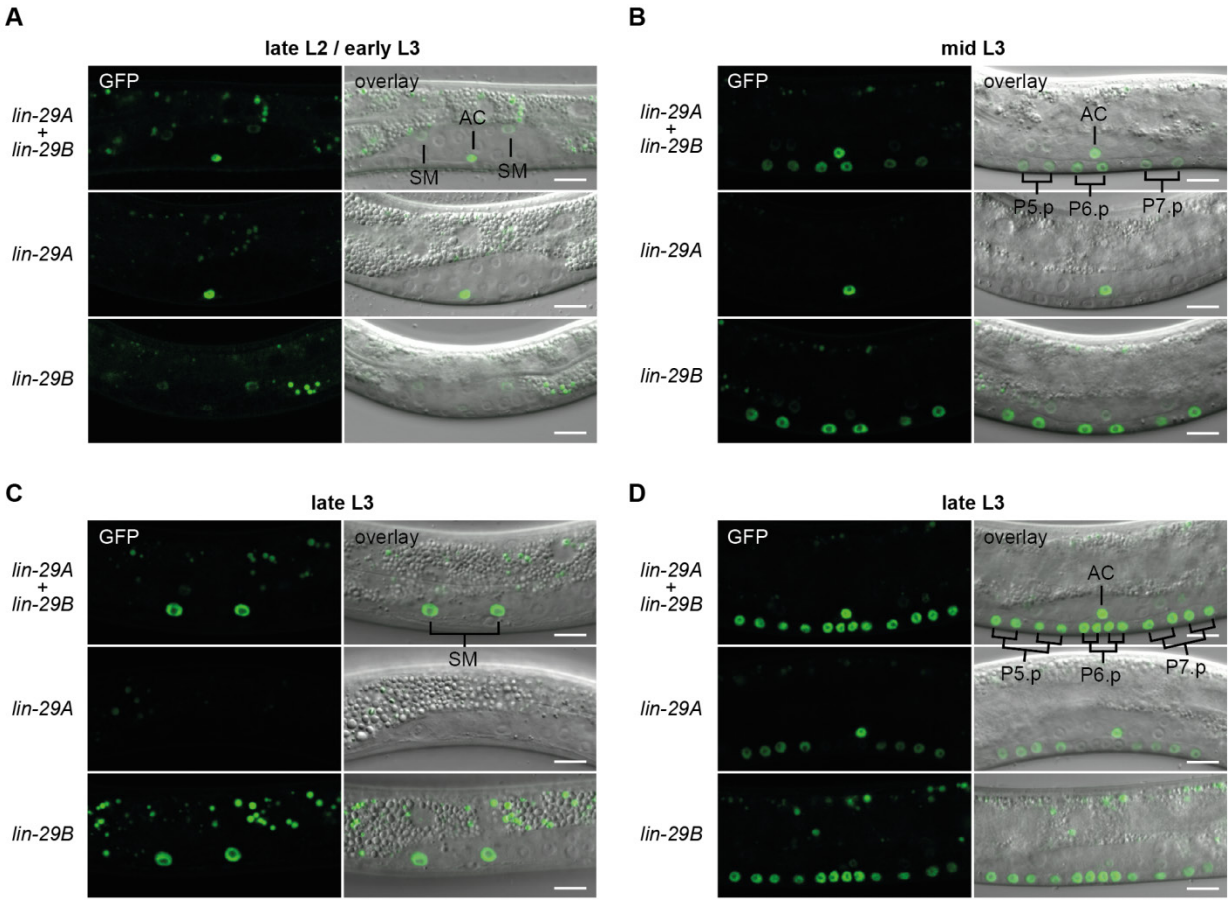


Figure 15. Expression patterns of LIN-29A and LIN-29B in the vulval region
(A-D) Confocal images of endogenously tagged LIN-29 protein isoforms in the region of vulval and uterine development at the indicated developmental stages. Scale bars: 10 µm. Starting at the L2-to-L3 molt (A), LIN-29A is expressed in the anchor cell (AC), while LIN-29B is weakly expressed in the sex myoblasts (SMs). In mid L3 stage worms (B), the six daugthers of the VPCs P5.p-P7.p express LIN-29B. At the late L3 stage, LIN-29B is strongly expressed in the sex myoblast (SM) daughters (C) and all 12 granddaughters of the VPCs P5.p-P7.p (D), while LIN-29A specifically accumulates in the granddaughters of P5.p and P7.p, but not in those of P6.p.

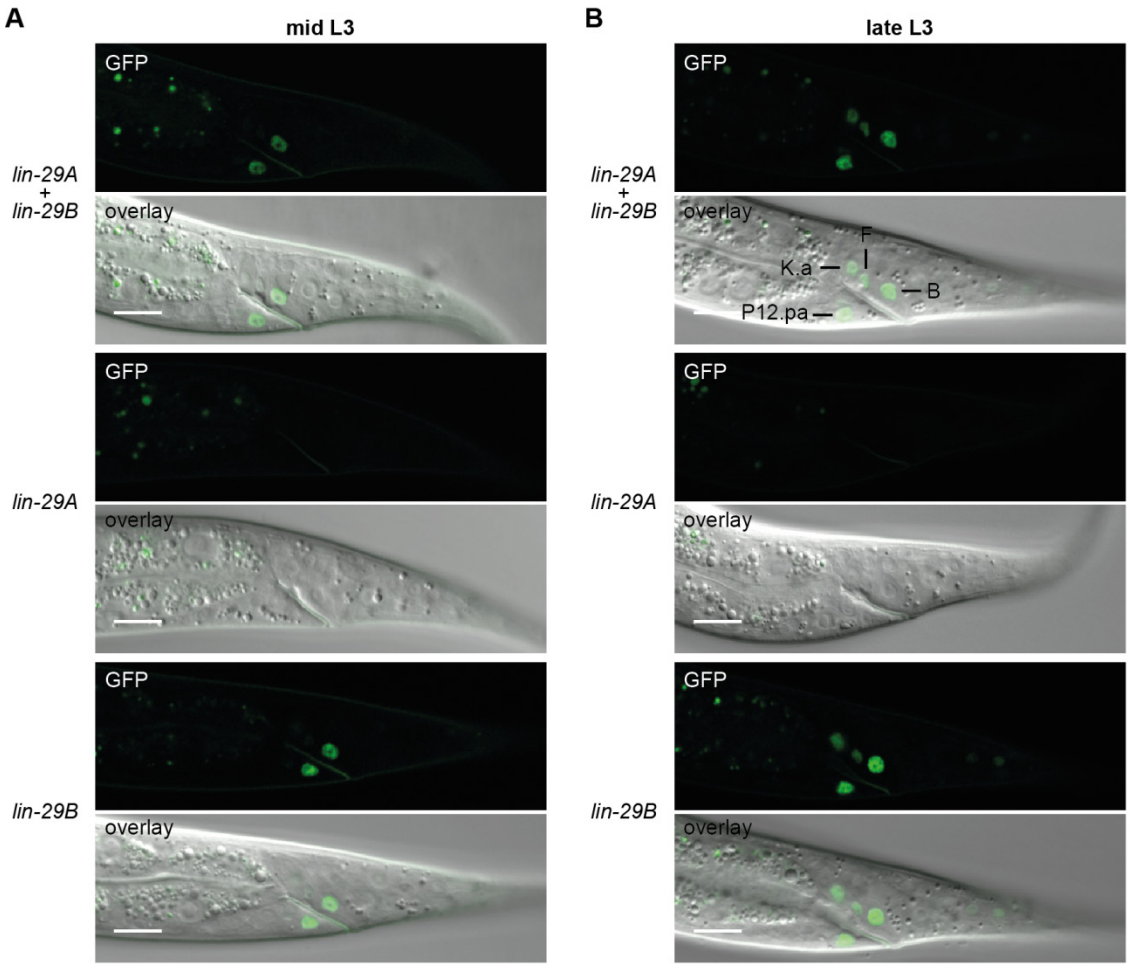


Figure 16. Expression patterns of LIN-29A and LIN-29B in the tail region
(A, B) Confocal images of endogenously tagged LIN-29 protein isoforms in the tail region of mid (A) or late (B) L3 stage animals. Scale bars: 10 µm. While LIN-29A cannot be detected, LIN-29B first accumulates in the two rectal cells B and P12.pa (A), before accumulating in the four additional rectal cells F, K.a, K' and U (B, the latter two are not visible in this focal plane).

2.7.5 A third LIN-29 isoform

Starting in L3 stage worms, we additionally detected a band with an intermediate size compared to LIN-29A and LIN-29B on the Western blot (Figure 17B). Although this band could correspond to a degradation product of LIN-29A, it is consistent with a third LIN-29 isoform that we can infer from our ribosome profiling data (Figure 17C). Visualization of RPF and RNA-seq wiggle files from L4 stage worms on a genome browser revealed RPF peaks on each annotated exon of lin-29A, including the 5'UTR exon, whereas there were no reads detected in the 3'UTR region. Unexpectedly, there was also a clear peak of RPF reads in a region within the long intron between *lin-29A* exons 4 and 5 (Figure 17C). These reads resided in a previously unannotated exon, referred to as exon 1c. This conclusion was based on RNA-seq reads

mapping to exon 1c and, more convincingly, mapping to an exon-exon junction to the downstream exon 5 (28 exon-exon junction spanning reads in total from both time course experiments, data not shown). When looking closer at the RPF peaks along this novel exon, we found an increase in RPF reads starting at an ATG that could serve as an initiation codon, as a very similar increase in RPF reads can be observed at the start codon for LIN-29A (data not shown). Translation initiation at the novel start codon yields a LIN-29 isoform with an N-terminal extension of 46 amino acids (MKSAEDSENYVLIDDDILSNCETENRRDLRKDMFEQKPDVGVLQQQ) compared to LIN-29B. This novel LIN-29 protein isoform, here referred to as LIN-29C, corresponds to a molecular weight of around 40.6 kDa (without GFP::3xFLAG tag), 5.4 kDa higher compared to LIN-29B and 9.6 kDa lower compared to LIN-29A, and thus fits the size of the detected band on the Western blot (Figure 17B). LIN-29C is likely not regulated by LIN41, as the RPF read coverage of this exon does not change to a similar extent as that of exons 1-4 in the *let-7(n2853)* mutant (Figure 17C).

Remarkably, an additional exon within the long intron between *lin-29A* exons 4 and 5 had been proposed before due to a sequenced cDNA clone and was named exon 1b, as it was thought to represent an alternative 5' end of the mRNA encoding for LIN-29B (Rougvie and Ambros, 1995). However, this exon is different from exon 1c (Figure 17C). We confirmed the existence of exon 1b in our RNA-seq data and found clear evidence for its splice junction to exon 5 (26 exon-exon junction spanning reads in total). Although we can detect RPF reads on exon 1b, mostly at 24 and 26 hours of development (data not shown), we think they probably reflect 5'UTR reads, as also seen for the *lin-29A* 5'UTR (Figure 17C). Consistent with this, translation initiation at the only ATG present in exon 1b would result in out-of-frame translation compared to the LIN-29 ORF. Therefore, we conclude that exon 1b is part of yet another mRNA isoform (referred to as *lin-29D* in Figure 17C), but does not result in production of yet another LIN-29 protein isoform.

From our RNA-seq data, it is unclear where the 5' ends of the mRNA isoforms *lin-29C* and *lin-29D* lie. SL1 trans-splice leader acceptor sites have so far only been identified for the beginning of exons 1 and 5, resulting in mRNA isoforms *lin-29A* and *lin-29B* (Figure 17C). For now, we annotated the 5' ends of *lin-29C* and *lin-29D* at potential splicing acceptor sites, at which we found very rare splicing events to upstream exons in our data, possibly representing splicing artefacts. In our analysis of isoform-specific expression patterns, we probably followed LIN-29C expression as part of the reported LIN-29B expression.

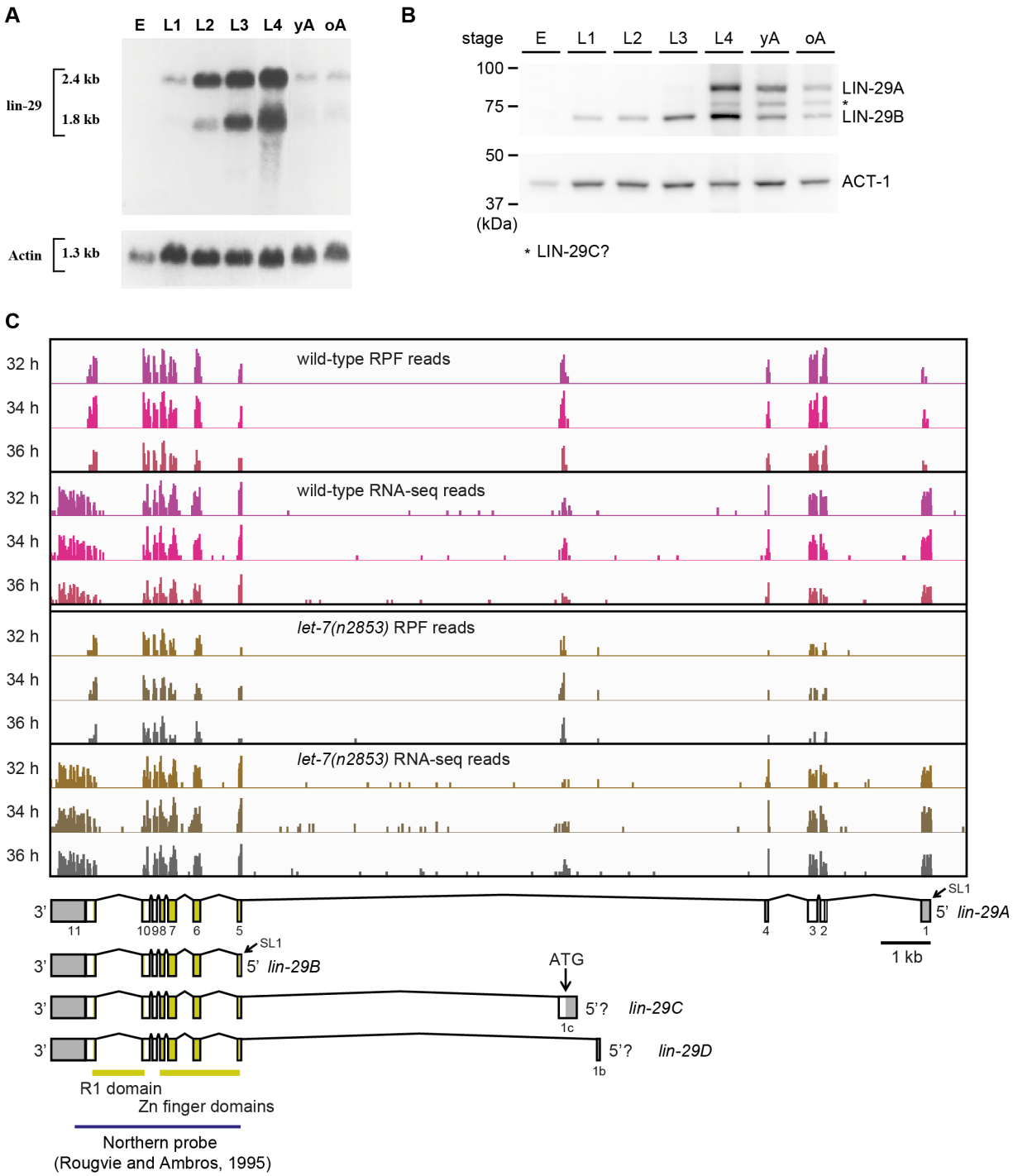


Figure 17. The dynamic expression of LIN-29A, LIN-29B and LIN-29C over development

(A) Northern blot analysis to detect the two *lin-29* mRNA isoforms at different stages in *C. elegans* development. Animals were harvested as embryos (E), as larvae of each of the four larval stages (L1-L4), and as young (yA) and older (oA) adults. Reproduced with permission from (Rougvie and Ambros, 1995).

(B) Western blot analysis to detect endogenous GFP::3xFLAG-tagged LIN-29A and LIN-29B proteins using an anti-FLAG antibody. ACT-1 was detected as a loading control. Animals were harvested at the same developmental stages as for the Northern blot (A). The asterisk marks the band that potentially corresponds to LIN-29C.

(C) Genome browser (IGV) screen shots visualizing accumulation of ribosome-protected fragment (RPF) or mRNA sequencing reads, processed to wiggle files, at 32, 34 and 36 hours of development at 25 °C.

2.8 LIN-29A and LIN-29B are regulated by different members of the heterochronic pathway

As described in the introduction, the heterochronic gene *hbl-1*, the homolog of the fly hunchback gene, acts partially redundantly with *lin-41*. This redundancy was observed at the genetic level (suppression of *lin-4* mutant phenotypes) (Abrahante et al., 2003), but also at the phenotypic level: Whereas single depletions of LIN41 and HBL-1 resulted in partially penetrant precocious phenotypes (see below), the depletion of both LIN41 and HBL-1 led to an almost complete execution of the L/A switch one stage too early, and some worms display precocious alae even two stages too early (Abrahante et al., 2003; Lin et al., 2003). However, because the direct targets of *hbl-1* and *lin-41* had not been identified, the nature of this redundancy remained completely unclear. Genetic experiments suggested that both *lin-41* and *hbl-1* somehow act through *lin-29*. This was because *lin-41; lin-29* and *hbl-1; lin-29* double mutants resembled *lin-29* single mutants, i.e. instead of showing precocious phenotypes, they repeated larval fates and never went through the L/A switch (Abrahante et al., 2003; Lin et al., 2003; Slack et al., 2000).

In the literature, both *lin-41* and *hbl-1* single mutants are sometimes superficially described as animals with seam cells that go through the epidermal L/A switch one stage too early. However, when looking closer at the results of previous studies (Abbott et al., 2005; Abrahante et al., 2003; Lin et al., 2003; Slack et al., 2000; Vadla et al., 2012), it becomes clear that these mutants display only partially penetrant and non-overlapping defects. In the absence of LIN41, precocious alae at the L3-to-L4 molt are observed in only about 50 % of animals (Slack et al., 2000), and these are described as small patches of well-formed alae, in contrast to those formed after co-depletion of HBL-1, described as full-length and strong alae (Abrahante et al., 2003). Precocious, often complete alae are also found on animals depleted of only HBL-1 with full penetrance, but these alae were described as weak structures (Abrahante et al., 2003). Moreover, these animals showed precocious and largely complete seam cell fusion at the L3-to-L4 molt, whereas the seam cells did not exit the cell cycle and continued to proliferate (Abrahante et al., 2003; Lin et al., 2003).

The phenotypes described for HBL-1-depleted animals at the L3-to-L4 molt were reminiscent of those we observed for the *lin-29(ΔA); mab-10(0)* double mutants one stage later (section 2.6). As in the latter mutants, LIN-29B, but not LIN-29A, is expressed, we speculated that depletion of *hbl-1* leads to precocious and specific expression of LIN-29B. Consistent with this idea, LIN-29 was reported to be precociously expressed in seam cells and not hyp7 in L3 stage *hbl-1(ve18)* mutants (Abrahante et al., 2003).

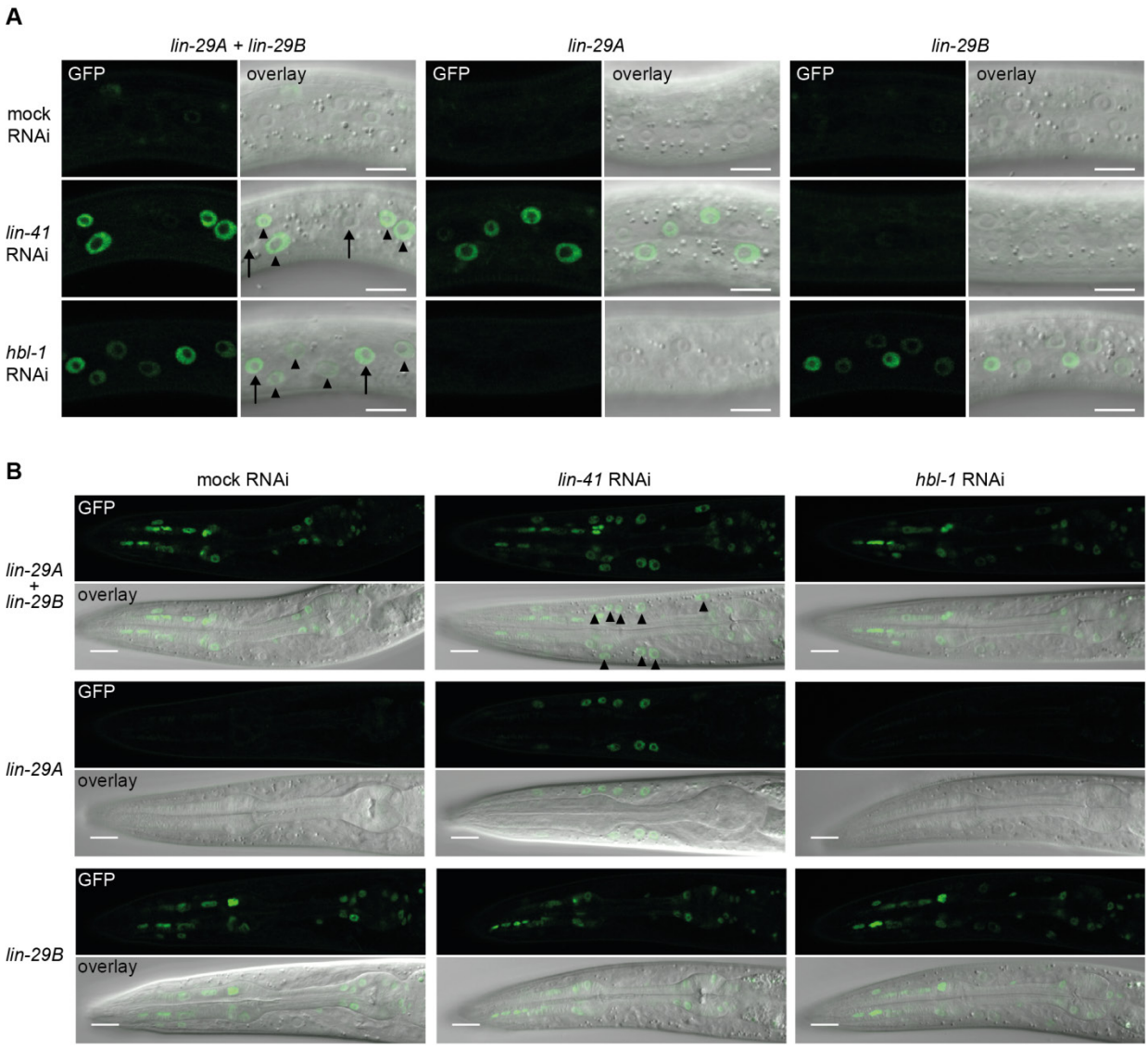


Figure 18. LIN-29A and LIN-29B expression are specifically regulated by LIN41 and HBL-1, respectively
(A, B) Confocal images of endogenously tagged LIN-29 protein isoforms in the epidermis (A) or pharynx (B). Animals were grown for 20h at 25 °C on *lin-41, hbl-1* or mock RNAi bacteria. Scale bars: 10 µm.
(A) Upon LIN41 knockdown, LIN-29A accumulates specifically in hyp7 (arrowheads) and not seam cells (arrows). Knockdown of HBL-1 results in specific accumulation of LIN-29B, mainly in seam cells, but also weakly in hyp7.
(B) The pharyngeal expression of LIN-29B is not changed upon HBL-1 or LIN41 knockdown. In this central focal plane, some nuclei of hyp7 (arrowheads) are visible that accumulate LIN-29A upon LIN41 knockdown.

To test this idea, we studied the expression patterns of LIN-29A and LIN-29B in the epidermis of early L3 stage worms harbouring a GFP::3xFLAG tag on different LIN-29 isoforms. Upon depletion of either LIN41 or HBL-1 through RNAi, we observed a precocious LIN-29 signal in the epidermis (Figure 18A), consistent with previous reports (Abrahante et al., 2003; Slack et al., 2000). However, the two expression patterns were completely different. Whereas we detected LIN-29 only in hyp7 cells in worms depleted of LIN41,

we observed LIN-29 accumulation mostly in seam cells in worms depleted of HBL-1. Strikingly, the precocious LIN-29 accumulation in absence of HBL-1 was specific to LIN-29B, with no precocious accumulation of LIN-29A (Figure 18A). On the contrary, depletion of LIN41 resulted in precocious accumulation of LIN-29A, but not LIN-29B (Figure 18A), confirming our previous results.

As described above, LIN-29B is expressed in many non-epidermal cells at the L2 and L3 larval stages. Therefore, we tested if depletion of HBL-1 would change the expression of LIN-29B in additional tissues. In the pharynx, where LIN-29B is already present in early larval stages of wild-type worms (Figure 14D), its expression did not change in absence of HBL-1 (Figure 18B). By contrast, we observed an increase in LIN-29B expression in two other regions: On the one hand, we observed precocious accumulation of LIN-29B in two rectal cells (Figure 19A), likely corresponding to the B and P12.pa cells that normally express LIN-29B only by the mid L3 stage (Figure 16A). On the other hand, we detected LIN-29B expression in the VPCs (Figure 19B). Importantly, at the time point of our analysis, we observed that the VPCs of HBL-1 depleted worms have divided at least once, but usually twice, revealing strong LIN-29B expression in the precocious VPC granddaughters. Indeed, mutants of *hbl-1* have previously been shown to go precociously through developmental events not only in the epidermis, but also in the vulva. Specifically, the P5.p-P7.p cells of HBL-1-depleted worms can divide one stage too early, in the extreme case resulting in the start of vulva morphogenesis by the L2-to-L3 molt (Abrahante et al., 2003). The LIN-29B expression pattern in the VPC granddaughter cells thus reflects the wild-type situation (Figure 15), but one stage too early. However, although in wild-type worms, the P5.p and P7.p granddaughters also express LIN-29A, this was not the case for the corresponding precocious granddaughters. Among other scenarios, this could suggest that *lin-29A* in these cells is repressed by LIN41.

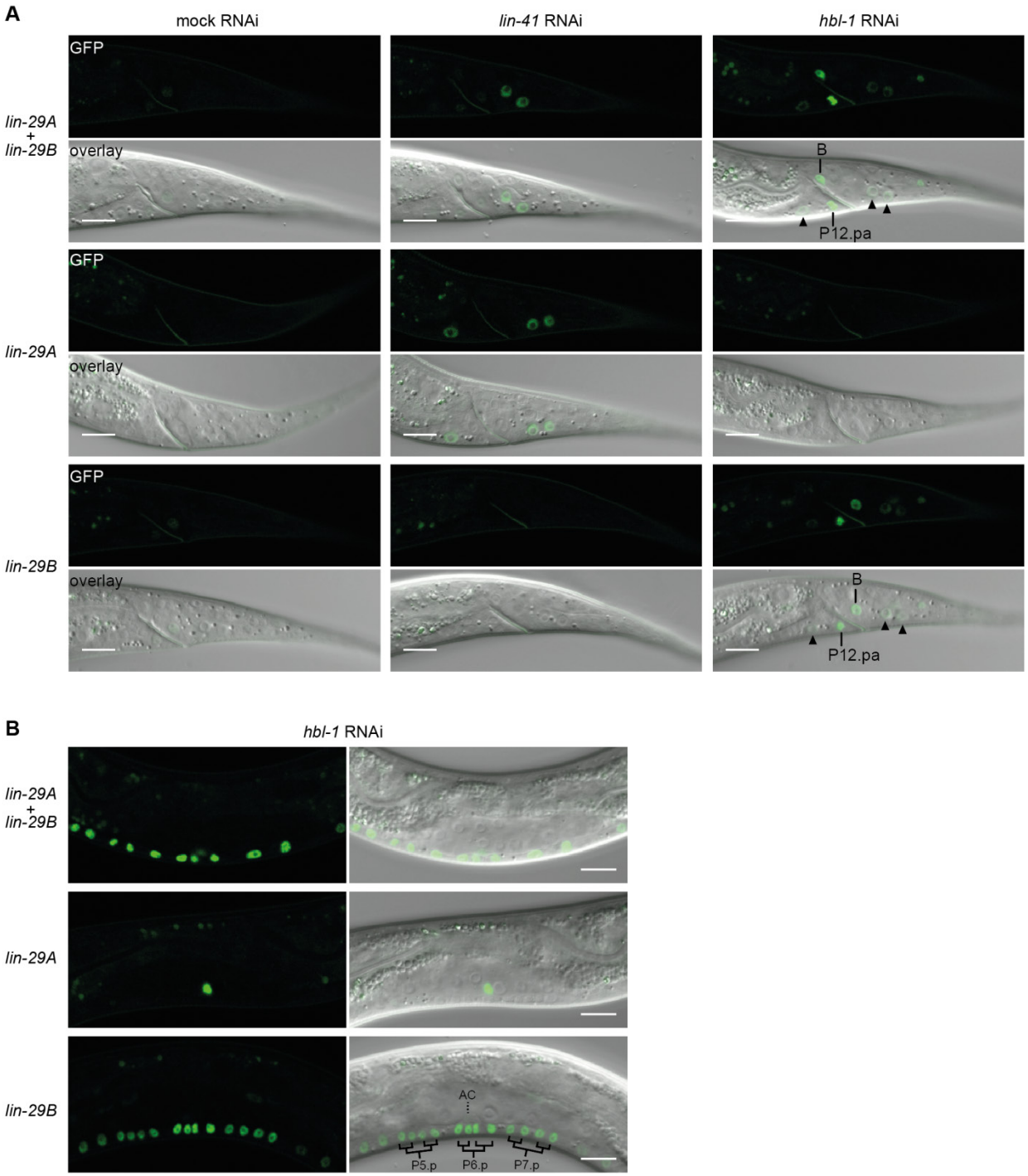


Figure 19. LIN-29B precociously accumulates in rectal and vulval cells in worms depleted of HBL-1
(A) Confocal images of endogenously tagged LIN-29 protein isoforms in the tail region. Animals were grown for 20h at 25 °C on *lin-41, hbl-1* or mock RNAi bacteria. LIN-29B specifically accumulates in rectal cells B and P12.pa upon HBL-1 depletion, while LIN-29A accumulation can be detected in some hyp7 nuclei (arrowheads) of the same focal plane. Scale bars: 10 µm.
(B) Confocal images of endogenously tagged LIN-29 protein isoforms in the vulva region of animals grown for 20h at 25 °C on *hbl-1* RNAi bacteria. The precociously dividing VPCs P5.p-P7.p accumulate LIN-29B, but not LIN-29A. Scale bars: 10 µm.

3 Discussion

3.1 The position-dependent dual activity of LIN41: unprecedented but not unique

Unlike most other RBPs, LIN41 can repress its targets by binding to both, the 3'UTR or the 5'UTR. Thereby, the binding location instructs the choice between two different post-transcriptional mechanisms: LIN41 elicits pure translational repression when bound to the 5'UTR and mRNA degradation when bound to the 3'UTR. Although other RBPs like Ago proteins (section 1.2.1) share the ability of inducing both mRNA degradation and translational silencing, LIN41 seems to be the first of its kind, as we do not know of any other RBP with the same binding site-dependent dual activity. Nevertheless, position-specific functions might constitute a more common, presently underappreciated feature of RBPs, since at least two other instances of regulatory RBPs with distinct functions at the 5'UTR and the 3'UTR have been described.

3.1.1 Iron regulatory proteins and sex-lethal elicit distinct mechanisms at the 5'UTR and 3'UTR

Firstly, the iron regulatory proteins 1 and 2 (IRP1 and IRP2) bind to RNA stem-loop structures in UTRs named iron responsive elements (IREs) to regulate expression of proteins involved in iron metabolism (Kühn, 2015; Volz, 2008). Binding of IRPs to mRNAs results in inhibition of translation initiation when IREs are located in the 5'UTRs, but results in protection from degradation when IREs are located in the 3'UTRs (Figure 20A). Mechanistically, to inhibit translation initiation, IRPs bind to a cap-proximal hairpin and block binding of the 43S preinitiation complex to eIF4F. This intuitive mechanism can be mimicked by tethering other high affinity RBPs near the cap, sterically blocking recruitment of the 43S complex (Gray and Hentze, 1994; Stripecke and Hentze, 1992). How 3'UTR-bound IRPs protect their target mRNAs from degradation has yet remained unknown. Through these two distinct mechanisms, IRP1 and IRP2 control cellular iron homeostasis, e.g. by regulating translation initiation of ferritin and ferroportin mRNAs, and by regulating degradation of the transferrin receptor 1 (TfR1) mRNA. Interestingly, an IRP-resistant splice variant of the ferroportin mRNA without the IRE in the 5'UTR is expressed in duodenum and erythroid cells (Cianetti et al., 2005; Zhang et al., 2009), reminiscent of the expression of LIN-29B in certain worm tissues.

Secondly, in a failsafe mechanism to ensure a complete block of translation, sex-lethal (SXL) binds to both the 3'UTR and the 5'UTR of the *male-specific lethal (msl)-2* mRNA in female flies (Bashaw and Baker, 1997; Kelley et al., 1997), eliciting distinct mechanisms of translational repression (Beckmann et al., 2005) (Figure 20B). On the 3'UTR, SXL inhibits the binding of the 43S preinitiation complex to the 5' end of the mRNA in a mechanism that is dependent on the RNA-binding protein UNR (upstream of N-ras) bound to adjacent 3'UTR binding sites (Abaza et al., 2006; Duncan et al., 2006; Hennig et al., 2014). On the other

hand, SXL bound to the 5'UTR stalls scanning 43S preinitiation complexes and enhances translation initiation at an upstream open reading frame (uORF), reducing translation from the main ORF (Beckmann et al., 2005; Medenbach et al., 2011). These mechanisms ensure that 43S preinitiation complexes escaping the 3'UTR mediated inhibition are hindered from translating the main ORF. The robust inhibition of MSL-2 activity, achieved by these elaborate mechanisms, is essential for the survival of female flies (Bashaw and Baker, 1997; Kelley et al., 1997). In male flies, where SXL is not functional, MSL-2 is produced and crucial for the assembly of the dosage compensation complex, which mediates a two-fold upregulation in transcription of the single X chromosome (Conrad and Akhtar, 2011).

3.1.2 Possible mechanisms of LIN41-mediated repression

The dual activities for IRP and SXL were discovered more than 25 and more than 10 years ago, respectively. Thus, the uncovered mechanistic details for their mechanisms of action are the result of many years of research. In this work, we discovered that LIN41 elicits different mechanisms on the 5'UTR and the 3'UTR (Figure 20C), opening up a new area of research for an RBP with position-dependent activities. From our data, we cannot make conclusions about the mechanism of 3'UTR-bound LIN41 to induce mRNA degradation on its targets. However, the ribosome profiling data provides hints to possible mechanisms of translational repression by 5'UTR-bound LIN41. Analysis of RPF read densities along a reading frame can help to distinguish different modes of translational repression. Inhibition of translation elongation or premature ribosome drop-off both result in less translating ribosomes towards the 3' end of the message and thus should reduce the RFP density at the 3'end to a larger extent than at the 5'end of the ORF. We observed the opposite for *lin-29A*, a larger decrease in RFP reads at the 5'end compared to the 3'end upon LIN41 binding (Aeschimann et al., 2017). However, this pattern reflects the differential regulation of *lin-29* isoforms rather than a mode of translational repression. When tracking RFP densities only along the exons of *lin-29A* (exons 1-4), or only along those shared between *lin-29A* and *lin-29B* (exons 5-11), LIN41 binding resulted in a homogenous decrease in RPF reads on all those exons. This argues against a regulation of translation elongation or ribosome drop-off and rather suggests inhibition of translation at the initiation step.

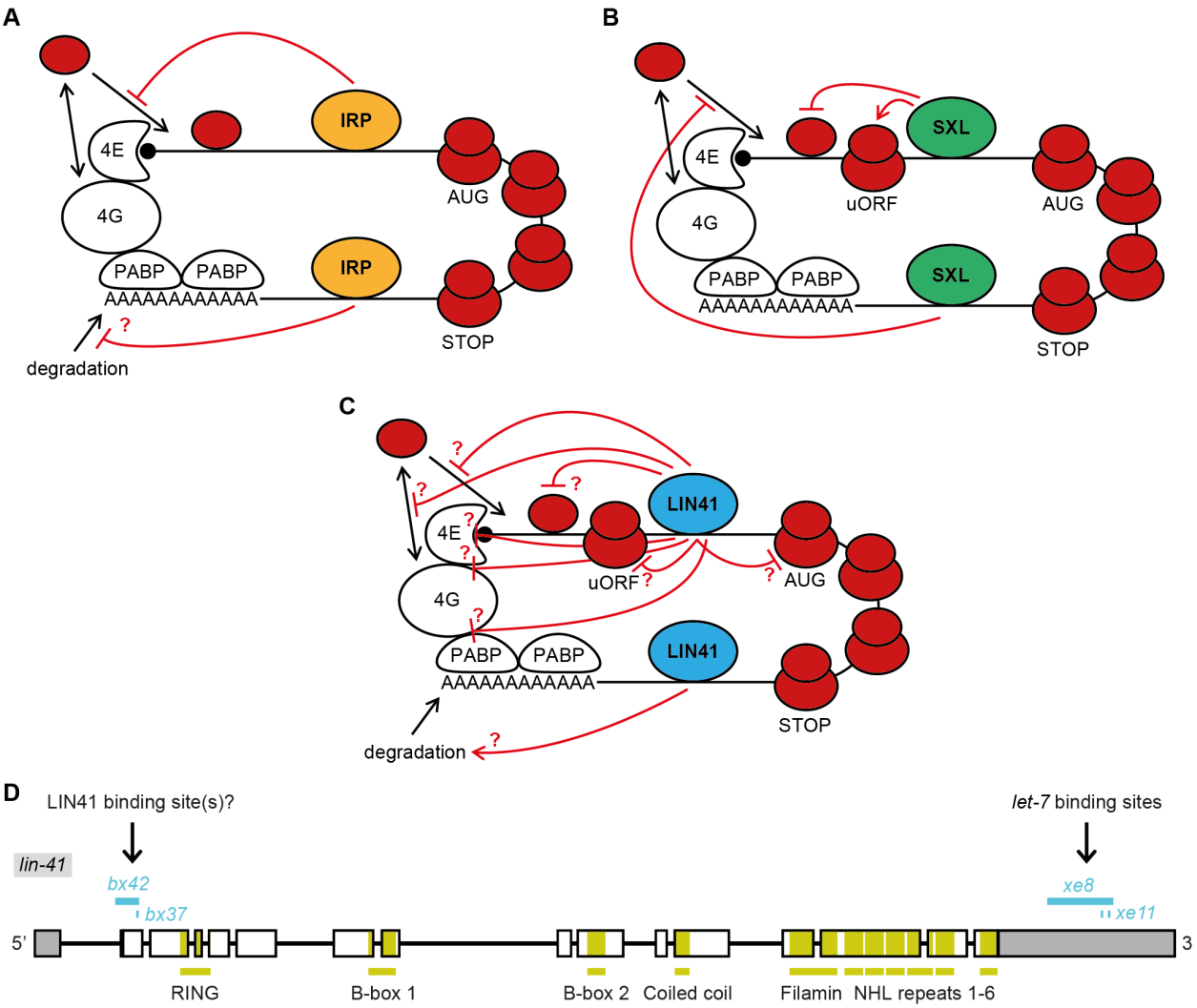


Figure 20. RBPs with distinct functions when bound to 5'UTRs and 3'UTRs, respectively.
(A) IRPs bound to the 5'UTR sterically inhibit recruitment of the 43S preinitiation complex to repress translation initiation, while IRPs bound to the 3'UTR protect mRNAs from degradation through an unknown mechanism
(B) SXL bound to the 5'UTR reduces translation initiation on the main ORF in two ways: It enhances translation initiation at an uORF and inhibits scanning of 43S complexes. SXL bound to the 3'UTR, together with its associated protein UNR (not drawn), inhibits recruitment of the 43S complex to the mRNA, also resulting in translational repression at the initiation step.
(C) LIN41 bound to the 5'UTR inhibits translation, most likely at the initiation step. The exact mechanism remains unclear and several possible, not mutually exclusive mechanisms are depicted. If LIN41 inhibits the binding of the 43S complex, it could do so by directly blocking its association or by interfering with protein-protein interactions necessary for efficient 43S complex recruitment. Those include interactions of eIF4E to the cap and interactions of eIF4G to eIF4E, PABP or eIF3. Alternatively, LIN41 could inhibit 43S complex scanning or recruitment of the 60S subunit at initiation codons, including those at uORFs. LIN41 bound to the 3'UTR enhances degradation of the target mRNA, employing a mechanism that remains to be determined.
(D) Illustration of different gain-of-function (gf) mutant alleles (in blue) of *lin-41*. All gf alleles reside in regions outside annotated domains (colored in yellow). The *bx42* allele is a 124 bp deletion starting in the first intron of *lin-41* and removing the first 88 bp of the *lin-41* ORF. The *bx37* allele is a G-to-A point mutation corresponding to a glycine-to-arginine transition at residue 34 of the *lin-41* ORF. The *xe8* and *xe11* alleles are explained in sections 2.2 and 2.5.

This mode of repression seems consistent with another observation: In the absence of LIN41, we detected RPFs on the non-coding exon 1 of *lin-29A*, a pure 5'UTR sequence lacking potential AUG start codons (Aeschimann et al., 2017). The nature of these reads, accumulating at lower numbers than those in coding exons, is unclear to us. According to the ribosome profiling protocol (Aeschimann et al., 2015), these 5'UTR reads should represent RPFs from assembled 80S ribosomes. Thus, it is possible that they originate from uORF translation initiating at codons other than AUG. Accordingly, their absence upon binding of LIN41 suggests that LIN41 inhibits any assembly of 80S ribosomes on the 5'UTR, excluding that LIN41 acts like SXL to increase translation initiation at an uORF. Therefore, we speculate that LIN41 either inhibits the recruitment of the 43S complex, 43S complex scanning or recruitment of the 60S subunit (Figure 20C).

The fact that transplantation of the *lin-29A* 5'UTR exon into an unregulated 3'UTR elicited mRNA degradation, whereas transplantation of a 3'UTR binding site from *mab-10* into a 5'UTR elicited translational inhibition (Aeschimann et al., 2017), is a strong argument that the location of the LIN41 binding site instructs the silencing mechanism. An alternative experiment to verify that binding of LIN41 to locations 5' or 3' of the ORF results in distinct repression mechanisms would be to tether LIN41 to these regions in a reporter construct. However, tethering experiments at the 5'UTR have been problematic in the past. Translational repression was not observed when SXL was tethered to the 5'UTR of a reporter (Grskovic et al., 2003), and proteins otherwise not known to be translational repressors have led to inhibition of translation when tethered to the 5'UTR (Gray and Hentze, 1994; Stripecke and Hentze, 1992).

3.2 Strategies to characterize the mechanisms of LIN41-mediated repression

3.2.1 Identification of LIN41 interaction partners

To obtain more insights into the mechanism of LIN41 as a translational repressor or as an activator of mRNA degradation, several different strategies can be pursued. One important venture is the identification of LIN41-interacting proteins. Based on our knowledge about other RBPs repressing translation or inducing mRNA degradation (section 1.2), it seems unlikely that LIN41 acts by itself. For acceleration of mRNA degradation, LIN41 is likely to directly or indirectly interact with factors of mRNA decay pathways, such as deadenylases, decapping factors and exonucleases. Although translational inhibition could be induced through steric hindrance of 43S complex recruitment or scanning, it could also involve interactions to translational repressors such as 4E-BPs.

Hence, important hints for mechanistic details could be obtained by finding physical interaction partners of LIN41. The functional transgenic FLAG::GFP::LIN41 protein we used for RNA-IPs ((Aeschimann et al., 2017) and section 2.4) could be used to perform IPs coupled to mass spectrometry for identification of co-immunoprecipitated proteins, both dependent and independent on the presence of RNA. Alternatively, genetic screens could be performed to identify genes that are necessary for reporter-based LIN41-mediated translational repression and/or mRNA degradation. To do so, the reporters with LIN41 binding sites on the 5'UTR or the 3'UTR (Aeschimann et al., 2017) could be used to screen for mutants or RNAi conditions which de-silence the reporter, similar to a knockdown of LIN41 (Aeschimann et al., 2017). An advantage of such screens is that they should reveal only LIN41 interactors involved in the repressive activity of LIN41 and not any other LIN41-associated proteins that may also be detected in a co-IP. On the other hand, these genetic screens could also yield hits with more indirect effects on the GFP reporter, for example mutations or knockdown conditions that decrease transcription of the *lin-41* gene. Therefore, the combination of both approaches could be a successful strategy to find the co-factors of LIN41 relevant for achieving its functions as a post-transcriptional gene repressor.

Since LIN41 is likely to assemble a complex of proteins on its target mRNAs, it would be interesting to identify the proteins that are bound to LIN41 target UTRs. This could be achieved using streptavidin-coated magnetic beads to pull-down biotinylated RNAs, such as the *lin-29A* 5'UTR exon RNA sequence, after incubation with a *C. elegans* lysate. Bound proteins identified by mass spectrometry could then be compared after incubation with wild-type and *lin-41* mutant lysates, respectively. An interesting alternative could be the identification of RBPs that bind to the LIN41 target mRNAs *in vivo*, for example by adapting a strategy that has recently been successful in identifying the RBPs associated with the long non-coding RNA *Xist* (McHugh et al., 2015). A main obstacle in this approach may be the hundreds of RBPs

associating with the mRNAs, independent of LIN41-dependent repression. However, this may be less problematic for the *lin-29A* mRNA, since in the presence of LIN41, *lin-29A* is devoid of translating ribosomes that likely constitute the main contaminants in such co-IPs. As another idea, the complexity of the analysed mRNPs could be reduced by using reporter minigenes with LIN41 binding sites, harbouring minimal ORFs and UTRs.

Except for the NHL repeat domain that mediates RNA-binding (Loedige et al., 2013; Loedige et al., 2015), it has remained unclear which domains of LIN41 are important for translational repression or mRNA degradation of its target mRNAs. Tethering experiments with the human LIN41 homolog revealed that B-box 2 and/or coiled-coil domains are important for target repression (Loedige et al., 2013), but a more detailed analysis of different domain mutants is needed to more clearly define the effector domains for target repression. Isolated effector domains could be used for protein co-IPs instead of the whole protein, potentially increasing transgenic expression and specificity of associated proteins, an approach that has for instance proven successful for identification of GW182-interacting proteins (Chekulaeva et al., 2011). The LIN41 coiled-coil domain may be a good candidate for such an approach, since coiled-coil domains have been implicated in formation of homo- or heterodimers, or even larger protein complexes (Burkhard et al., 2001). Thus, it is also possible that LIN41 needs to homodimerize in order to silence its targets. A need for two RNA-bound LIN41 molecules for repression could be an explanation to why it was not possible to find a shorter stretch of the *lin-29A* 5'UTR exon that still mediates translational repression (Aeschimann et al., 2017).

3.2.2 Narrowing down the post-transcriptional repression mechanisms

Independent of the identification of LIN41-interacting proteins, the repression mechanisms themselves could be better characterized. The lower mRNA abundance of *mab-10*, *mab-3* and *dmd-3* in the presence of LIN41 is likely due to enhanced mRNA degradation, but to explicitly test this, their mRNA half-lives would have to be measured in the presence and absence of LIN41. Moreover, it could be tested if deadenylation is involved in the repression mechanisms, by measuring poly(A) tail lengths of LIN41 target mRNAs in presence and absence of LIN41, e.g. with ePAT assays (Jänicke et al., 2012). Alternatively, poly(A) tails could be measured globally by high-throughput sequencing methods such as TAIL-seq or PAL-seq (Chang et al., 2014; Subtelny et al., 2014), the former technique also allowing for detection of potential degradation intermediates containing untemplated nucleotides other than As at the 3'end (Chang et al., 2014).

To narrow down the mechanism of translational inhibition through LIN41, ribosomal association of the *lin-29A* mRNA could be analysed with high resolution sucrose gradients, similar to experiments performed with the SXL-bound *msl-2* mRNA (Beckmann et al., 2005; Gebauer et al., 2003). Thereby, the *lin-29A* mRNA abundance could be measured by RT-qPCR in the different fractions of the gradient, to compare its distribution in presence and absence of LIN41. The addition of different chemical inhibitors of translation would then allow to determine if LIN41 affects 80S or 48S (43S complex at the start codon) complex formation. These experiments would help to verify translational inhibition at the initiation step and to distinguish between mechanisms blocking 43S complex recruitment or scanning and mechanisms to inhibit 60S subunit joining (Figure 20C). To distinguish between inhibition of 43S complex recruitment and inhibition of scanning, 80S or 48S complex formation could be analyzed for a reporter construct with an AUG start codon inserted into the beginning of the *lin-29A* 5'UTR. If LIN41 on its natural target inhibits 43S complex scanning in the proximity of its binding site, it may not affect 48S and 80S complex formation on the introduced upstream AUG. On the other hand, if LIN41 acts by inhibiting recruitment of the 43S complex, the reporter mRNA would be unable to form 48S or 80S complexes.

3.3 How does LIN41 regulate its own activity?

Numerous different *lin-41* alleles have been isolated in many studies (Del Rio-Albrechtsen et al., 2006; Ecsedi et al., 2015; Slack et al., 2000; Spike et al., 2014). Of those, there are two kinds of gf alleles: The *xe8* and *xe11* alleles located in the 3'UTR, which disrupt the binding to *let-7* (Ecsedi et al., 2015), and the *bx37* and *bx42* alleles located in the first exon of the *lin-41* coding sequence, where no domain has been annotated (Figure 20D). The two *bx* alleles cause identical phenotypes (section 2.5) and are thought to increase LIN41 levels due to the same mechanism, namely a failure in LIN41 autoregulation (Del Rio-Albrechtsen et al., 2006). LIN41 was proposed to autoregulate its activity, because addition of a wild-type *lin-41* copy to *lin-41*(*bx37)* gf mutant animals suppressed the gf phenotype, instead of making it more severe (Del Rio-Albrechtsen et al., 2006). This suggested that LIN41 can downregulate its own activity, and that this downregulation cannot occur in the *bx37* mutant background. Hence, LIN41 was proposed to repress its own activity through autoubiquitylation, an activity that would be disrupted in LIN41 proteins in the two *bx* mutant backgrounds, thereby stabilizing the pool of LIN41 proteins. Supplementary expression of wild-type LIN41 would lead to degradation of both wild-type and *bx37*-mutant LIN41 proteins, explaining the gf mutant suppression (Del Rio-Albrechtsen et al., 2006).

Our data (section 2.4), together with the finding that the RING finger domain of *C. elegans* LIN41 lacks an otherwise conserved proline necessary for E3 ubiquitin ligase activity (Tocchini et al., 2014), suggest a

different model for LIN41 autoregulation: LIN41 binds to its own mRNA to downregulate LIN41 protein production, through translational repression or mRNA degradation. Thus, an alternative explanation for the gf phenotypes of *bx37* and *bx42* is that the changes at the RNA level cause a gf phenotype, by affecting the binding site of LIN41 for repressing its own expression (Figure 20D). However, in this scenario, addition of a wild-type *lin-41* copy should not interfere with the gf phenotype, as the wild-type protein should not be able to bind to the mutant *lin-41* mRNAs. Nevertheless, if the two *bx* mutant alleles indeed disrupt the LIN41 binding site, this binding site is likely to reside in the coding region, where the two *bx* alleles map to, or in the neighboring 5'UTR, but not in the very distant 3'UTR.

If LIN41 indeed regulates its expression by binding to its own mRNA can only be determined with additional experiments. One approach would be to find the binding site(s) for LIN41 on its own mRNA, by assaying reporters containing the *lin-41* 5'UTR, 3'UTR, or the part of the coding sequence where the *bx37* and *bx42* alleles map to. An alternative strategy would be to directly test if the *bx* gf phenotypes are caused by changes at the mRNA level or at the protein level. To do so, silent mutations could be introduced by CRISPR-Cas9 into the DNA stretch to which the *bx37* and *bx42* mutations map, likely changing the RNA structure in this region but not the translated amino acid sequence. A resulting gf phenotype would be good evidence that a change in the RNA sequence in this region rather than a change in the protein can cause LIN41 overexpression. If in addition, a FLAG-tag is introduced to the N-terminus of this *lin-41* mutant, the resulting FLAG::LIN41 protein could be used in IPs to measure whether it normally binds to targets such as *lin-29* and *mab-10*, but not to its own mRNA.

3.4 The *let-7*-LIN41 module coordinates timing of several developmental events

3.4.1 LIN41 dys-regulation explains all known let-7 *phenotypes in C. elegans*

Although miRNAs are predicted to target numerous mRNAs and are thus commonly thought to function by regulation of many targets, at least the miRNAs acting in the heterochronic pathway seem to achieve their main functions by regulating one key target. Genetic experiments (section 1.4) have suggested that *lin-4* acts through its key target *lin-14*, the *let-7* sisters act through *hbl-1*, and this work, together with a previous study of our group (Ecsedi et al., 2015), establishes *lin-41* as the key target of *let-7*.

The miRNA *let-7* is a master regulator of the larval-to-adult transition, regulating and coordinating the timing of several different developmental events to ensure these do not occur too early or too late. These events include exit from the molting cycle, cessation of seam cell self-renewal and morphogenesis of the sexual organs of both the hermaphrodite and the male worm. At least the latter two of these *let-7* functions are achieved by regulation of a single target, the RNA-binding protein LIN41. This strong statement is possible thanks to genome editing experiments with the CRISPR-Cas9 system, which have made it possible to specifically rescue the *let-7* mediated silencing of *lin-41* in an otherwise *let-7(lf)* mutant background (Ecsedi et al., 2015). This very specific rescue completely abrogated the defects in vulval morphogenesis (Ecsedi et al., 2015), as well as male tail retraction and exit from seam cell self-renewal (sections 2.5 and 2.6). Although not examined so far, it is likely that these "rescued" animals also normally exit from the molting cycle. Supernumerary molts are observed in null mutants of *let-7*, *mab-10* and *lin-29* (Ambros and Horvitz, 1984; Harris and Horvitz, 2011; Reinhart et al., 2000), consistent with the *let-7*-LIN41 module regulating molting through LIN-29A and MAB-10. Strikingly, the decline in the capacity of neuronal regeneration as worms age was also shown to be dependent on *let-7* downregulating LIN41 (Zou et al., 2013), further emphasizing that completely different phenotypes of *let-7* mutant worms seem to depend on elevated LIN41 levels. In conclusion, the major, if not exclusive function of *let-7* consists of making sure that LIN41 is downregulated at the correct developmental time.

3.4.2 LIN41 acts through distinct targets to coordinate developmental events

Before *let-7* is upregulated, elevated LIN41 levels ensure that the epidermis, the vulva and the male tail keep their larval identity and do not start any gene expression programs to differentiate into adult tissues. As soon as *let-7* downregulates LIN41, all these tissues start their transformations into adult tissues, possibly at the same time. The knowledge gained in this thesis to understand the underlying mechanisms of these LIN41-mediated functions are summarized in Figure 21 and described below:

In early larval stages, LIN41 is produced and mediates position-dependent repression mechanisms on its mRNA targets. Thereby, LIN41 binds to the 5'UTR of *lin-29A* to inhibit translation and to the 3'UTR of *mab-10, mab-3*, *dmd-3* to induce degradation of the mRNA. Because LIN-29A and MAB-10 are not expressed, seam cells keep their self-renewal properties and proliferate, and certain, yet undefined processes in the vulva-uterine region do not take place yet. At the same time, the absence of MAB-3 and DMD-3 ensures that the male tail tip does not retract. In the last larval stage, LIN41 is targeted by *let-7* and thus not produced anymore. This in turn leads to the expression of the three transcription factors MAB-3, DMD-3 and LIN-29A, and of the LIN-29 cofactor MAB-10. Both LIN-29 isoforms can associate with MAB-10, and by binding to the promoters of their target genes, they either activate or inhibit transcription. The output of this change in the transcriptional program causes the seam cells to exit the cell cycle and stop the self-renewal program. In the vulva, the transcriptional change induces certain processes, yet to be identified, that are crucial for survival. The two DM domain transcription factors MAB-3 and DMD-3 bind to promoters of their target genes and act at least partially redundantly to signal the tail tip cells to start the retraction program. Although not examined in this thesis, this pathway could also regulate the cell fusion of the epidermal hyp8-hyp11 cells that specifically occurs in male animals and is completely absent in *mab-3;dmd-3* double mutants (Mason et al., 2008). In the model in Figure 21, the depicted LIN-29:MAB-10 complex, as well as MAB-3 and DMD-3, can activate or repress transcription on target genes. This is inferred from mammalian DM domain transcription factors, which can perform both these functions (Kopp, 2012), and from mammalian homologs of MAB-10, which were shown to act as coactivators or corepressors for distinct target genes (Sevetson et al., 2000; Svaren et al., 1998).

3.4.3 Genetic experiments to verify the proposed model: Does LIN41-mediated silencing of its identified targets cause the observed phenotypes?

Whereas there is convincing evidence that *let-7* acts through LIN41 as its sole key target to control the processes depicted in Figure 21 (section 3.4.1), we cannot be completely certain that LIN41 exclusively acts through *mab-10* and *lin-29A* to control seam cell self-renewal and through *mab-3* and *dmd-3* to control male tail tip retraction [The roles of *mab-10* and *lin-29A* in vulval morphogenesis are discussed in section 3.4.4]. From experiments with the null mutants of the LIN41 target genes, where we observed identical phenotypes to those seen with *lin-41(xe8)* mutants, we can conclude that it is sufficient to silence *lin-29A* and *mab-10* to cause overproliferation of seam cells (section 2.6) and *mab-3* and *dmd-3* to completely inhibit tail tip retraction (Mason et al., 2008). However, we have not formally shown that LIN41 needs to silence these targets in order to control seam cell proliferation and male tail retraction.

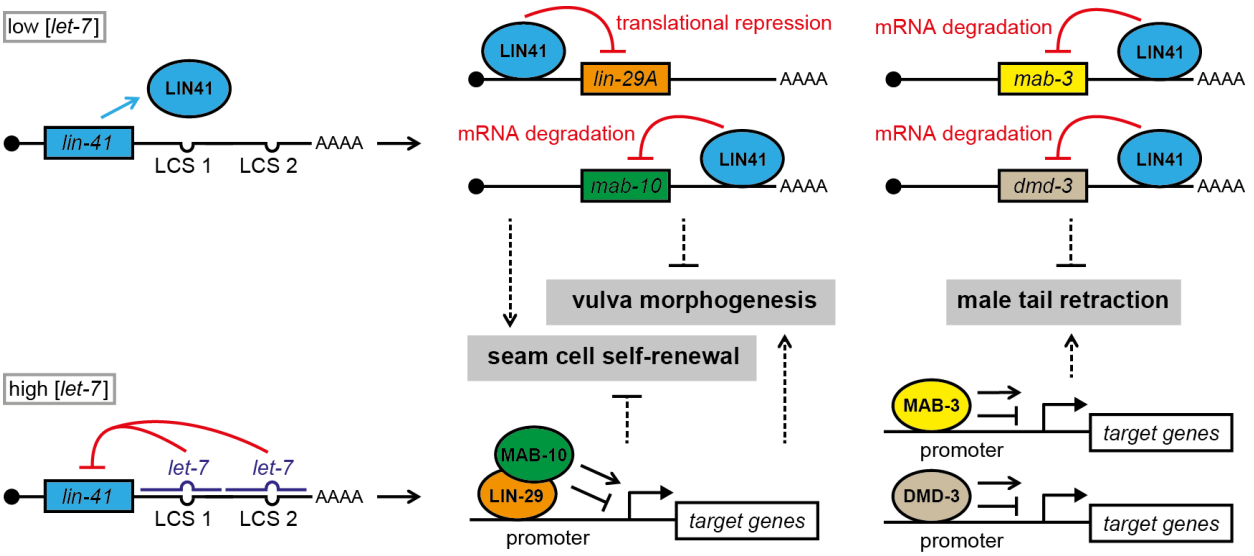


Figure 21. The *let-7*-LIN41 module coordinates different developmental events through distinct targets.
Model of the pathways downstream of the *let-7* miRNA that regulate seam cell exit from the cell cycle, proper vulva morphogenesis and retraction of the male tail tip. See main text for details.

To do so, it would be necessary to disrupt the binding of LIN41 to each individual target and observe the phenotypic consequences. Such experiments have become feasible due to the establishment of CRISPR-Cas9 in our lab (Katic and Grosshans, 2013; Katic et al., 2015) and may even be straightforward thanks to a recently published strategy for selection of CRISPR-Cas9-induced recombination events (Dickinson et al., 2015). If the model in Figure 21 is complete, we would expect that mutation of the LIN41 binding site in the 5'UTR of *lin-29A* (the hypothetical allele is here referred to as *lin-29A(ΔL41BS)*), combined with mutation of the binding site in the 3'UTR of *mab-10* (*mab-10(ΔL41BS)*), would cause epidermal phenotypes identical to those of *lin-41* null mutant animals. These phenotypes include premature alae secretion at the L3-to-L4 molt and, possibly, a premature termination of seam cell proliferation (Slack et al., 2000). The latter has been proposed for *lin-41* null mutants (Slack et al., 2000), but, to our knowledge, not closely examined. Accordingly, we would expect that mutation of the LIN41 binding sites in the 3'UTRs of *mab-3* and *dmd-3* (*mab-3(ΔL41BS)*; *dmd-3(ΔL41BS)* double mutant) elicits a premature retraction of the male tail tip cells, resulting in over-retracted adult male tails as seen for *lin-41(rf)* mutants (Del Rio-Albrechtsen et al., 2006). Furthermore, *lin-29A(ΔL41BS); mab-10(ΔL41BS)* and *mab-3(ΔL41BS)*; *dmd-3(ΔL41BS)* should be epistatic to *lin-41(xe8),* i.e. the phenotypes of triple mutants (the pairs of *ΔL41BS* alleles combined with *lin-41(xe8)*) should resemble those of the binding site double mutants and not those of *lin-41(xe8)*. Such epistasis would provide evidence that the LIN41 overexpression phenotypes can be suppressed by removing its relevant target sites, suggesting that *lin-41* indeed acts through our proposed targets.

If the expected phenotypes for *ΔL41BS* mutant combinations are observed, genetic rescue experiments could be designed, in order to test if restored LIN41-mediated downregulation can suppress these phenotypes. For this purpose, the LIN41 binding site of one target UTR could be introduced into another target UTR. In the example for studying male tail retraction, *ΔL41BS* alleles could be obtained by replacing the endogenous 3'UTRs of *mab-3* and *dmd-3* by the unregulated *unc-54* 3'UTR. In addition, the 3'UTRs of *mab-3* and *dmd-3* could be replaced by an *unc-54* 3'UTR containing an insert with a LIN41 binding site from the *lin-29* 5'UTR or *mab-10* 3'UTR. If the resulting mutant males normally retract their tails, whereas they do not in *mab-3(ΔL41BS)*; *dmd-3(ΔL41BS)* mutants, it could be concluded that LIN41-mediated male tail retraction is undoubtedly and exclusively due to its repression of *mab-3* and *dmd-3*.

3.4.4 Why do elevated LIN41 levels cause bursting?

The results of section 2.5 suggest that the lethal vulval rupturing observed upon a failure to downregulate LIN41 in late larval stages could be due to a change in activity of the transcription factor LIN-29. However, unlike the seam cell overproliferation phenotype, the fully penetrant vulva bursting phenotype of *lin-41(xe8)* mutants cannot be recapitulated by null mutations in *lin-29A* and *mab-10*. Although it remains possible that this incomplete penetrance reflects the involvement of additional LIN41 targets or the role of LIN41 as structural protein in this process, several indications hint towards a different model, namely that the bursting is caused by partial loss of LIN-29 activity: First, depletion of LIN-29 by RNAi causes bursting in more than 50 % of worms, compared to less than 2 % burst worms when LIN-29 is fully depleted in *lin-29(0)* mutants. Second, the *lin-29(rf)* allele *lin-29(n546)* causes bursting with much higher penetrance compared to *lin-29(0)* mutants. Third, the activity of LIN-29B alone in *lin-29(ΔA); mab-10(0)* mutants results in about 30 % burst worms, a much higher frequency compared to *lin-29(0)* mutant populations.

Different, not mutually exclusive explanations for these observations are possible, including the following: i) Bursting occurs only in a certain range of LIN-29 activity that is perfectly reached in *let-7(n2853)* or *lin-41(xe8)* worms. This would suggest that LIN41 does not completely silence LIN-29A and/or MAB-10, because *lin-29(ΔA); mab-10(0)* mutants burst with only 30 % frequency. ii) Worms burst when LIN-29 can be active in earlier larval stages (although to a modest extent and only in certain tissues (section 2.7)), but fails to increase its activity in later stages due to the persistence of high LIN41 levels. iii) Worms burst when LIN-29 activity in the L4 stage is silenced only in certain tissues but not others.

The last possibility is arguably the most plausible reason for vulval rupturing. Before LIN41 is downregulated, LIN-29A is highly expressed in the anchor cell and also modestly expressed in VPCs

(section 2.7). This suggests that LIN-29A cannot be repressed by LIN41 in these vulval cells, either because LIN41 is not expressed in these cells or because *lin-29A* somehow escapes the LIN41-mediated repression. Assuming that, accordingly, LIN41 in *lin-41(xe8)* L4 stage worms efficiently silences LIN-29A expression in all tissues except for some vulval cells, this could create an imbalance in developmental progression of different cells that eventually causes the bursting.

To test this model, the LIN-29A expression pattern should be studied in *let-7(n2853)* or *lin-41(xe8)* mutant animals, to identify the cells in which *lin-29A* silencing does not occur in late L4 stage worms. Subsequently, if possible, LIN-29A could be expressed in the identified (vulval?) cells from a transgene driven by a tissue-specific promoter, in a *lin-29(ΔA); mab-10(0)* mutant background. If this transgenic expression of LIN-29A causes bursting with full penetrance, a tissue-specific *lin-29A* silencing as the cause for bursting would be likely. However, tissue-specific *mab-10* silencing could also contribute to vulval rupturing and would have to be tested in a different experimental setup.

3.4.5 Why is lin-29A repressed on the translational level?

Translational repression is recognized as a post-transcriptional mechanism that allows rapid gene expression changes, well suited for different cellular stress responses (Szostak and Gebauer, 2013). A classic example is the translational repression of the *Saccharomyces cerevisiae* transcription factor GCN4, relieved upon amino acid deprivation, allowing transcriptional activation of genes involved in amino acid biosynthesis (Hinnebusch, 2005). However, a rapid upregulation of LIN-29A is probably not needed for its role in the larval-to-adult developmental transition, raising the question of why LIN41 represses *lin-29A* purely on the level of translation.

As one possible scenario, this mode of regulation could have evolved with the sole purpose of regulating only the LIN-29A isoform. As this can be achieved by binding to the *lin-29A* 5'UTR, translational repression as a silencing mechanisms could have just evolved by chance. Nonetheless, the repression on the level of translation could also represent an advantage over other repression mechanisms. Interestingly, LIN-29A is not the only protein with a crucial function in development that is regulated on the translational level. For example, several mRNAs critical for embryonic patterning in flies are translationally controlled (Johnstone and Lasko, 2001). As another example, SXL represses expression of the *msl-2* mRNA in female flies through translational inhibition (section 3.1), a mechanism that completely silences *msl-2* expression and is essential for the survival of female flies (Conrad and Akhtar, 2011). Translational repression could thus serve as a very robust silencing mechanism, to ensure that the target protein is not produced at all. At the moment, translational control serving as a well-suited mechanism for developmental on/off

switches is a very speculative model. However, it would be interesting to know if, compared to translational silencing, transcriptional control and control of mRNA turnover are less complete, i.e. more prone to transcripts escaping the silencing mechanisms.

At least in the case of LIN-29A, translational inhibition seems very tight. As long as LIN41 levels are high, no translational activity can be detected for *lin-29A*, despite high *lin-29A* mRNA levels ((Aeschimann et al., 2017), Figure 4B (18-22 h), a normalized $\log_2$ read count of 3 is equivalent to zero RPF reads). Consistently, no signal can be detected for endogenously tagged GFP::LIN-29A in the epidermis of L3 stage wild-type worms (section 2.8).

3.5 The L/A switch likely consists of independently regulated events

The L/A switch was described by Victor Ambros in 1989 as a developmental transition with four different easily detectable events (Ambros, 1989): Fusion of the seam cells, a discontinuation of seam cell divisions and of the molting cycle, and synthesis of alae. The L/A switch was considered as one entity, because all four events were regulated by a single gene, *lin-29*. Our analysis of the two different LIN-29 isoforms (sections 2.6-2.8) suggests that this old concept should be revised, and that the analysed activity of *lin-29* was the sum of the activities of two *lin-29* isoforms with different expression patterns and functions.

3.5.1 LIN-29A and LIN-29B have different roles during the L/A switch

Our study reveals some evident differences for the roles of LIN-29A and LIN-29B during the L/A switch: Whereas LIN-29B alone is sufficient for seam cell fusion, it needs the additional activity of MAB-10 to suppress seam cell proliferation and that of LIN-29A for wild-type alae formation (section 2.6). Thus, as outlined in Figure 22A in commonly used schemes to explain heterochronic mutants (section 1.4), mutation of *lin-29A* leads to weak alae structures (depicted in grey), mutation of *mab-10* leads to extra seam cell divisions, mutation of both *lin-29A* and *mab-10* results in weak alae and additional seam cell divisions, and only the mutation of both *lin-29* isoforms leads to a complete failure of all L/A switch events.

3.5.2 Some L/A switch events occur independently of other L/A switch events

Although incomplete due to the lack of a *lin-29B*-specific mutant, the analysis of the functions of LIN-29A, LIN-29B and MAB-10 clearly suggests that the four L/A switch events cannot be considered an entity and instead could be uncoupled events that are independently regulated. As illustrated in Figure 12, different L/A switch events are defective in different combinations of *lin-29(ΔA), lin-29(ΔAB)* and *mab-10(0)* mutants. Clear evidence against a model with four linked L/A switch events results from the analysis of seam cell fusion versus seam cell overproliferation (section 2.6). Seam cell fusion occurs normally in the mid L4 stage, in animals of different mutant backgrounds, independently of whether or not these animals subsequently go through extra seam cell nuclear divisions at the young adult stage. Exit from the cell cycle is thus clearly not elicited by cell fusion, and vice versa, cell fusion can occur independently from a cell cycle exit. Notably, seam cell fusion is not an irreversible event, as the syncytium can divide up into separate cells again after extra nuclear divisions occur (section 2.6).

Similarly, alae secretion is not concomitant with a stop in seam cell proliferation, as it occurs in both, animals which later do or do not display extra seam cell divisions. Therefore, next to the concept of the L/A switch, the concept of terminal differentiation also needs clarification, at least with respect to the heterochronic pathway. Terminal differentiation is commonly defined as a developmental switch which includes an exit from the cell cycle (Jones, 2007). However, in the heterochronic pathway literature, *C. elegans* seam cells are often considered terminal differentiated when alae structures are detected. This can lead to wrong conclusions, as, for example, *lin-29(ΔA); mab-10(0)* double mutants secrete (mostly complete) alae, but at the same time or even later in development, their seam cell nuclei divide again. The same is true for *hbl-1(rf)* mutants (Abrahante et al., 2003; Lin et al., 2003). In summary, the L/A switch, and thus also the terminal differentiation program, likely consists of separate, unlinked events. Therefore, alae structures should neither be used as a marker for a completed L/A switch, nor for terminally differentiated seam cells.

3.5.3 The impact of LIN-29A, LIN-29B and MAB-10 on gene expression

The data in this thesis contradicts the proposed model that LIN-29A and LIN-29B redundantly contribute to the same functions (Bettinger et al., 1997). A new model in which LIN-29A and LIN-29B have different roles in *C. elegans* development opens up numerous questions to be answered by future research. Next to obtaining a *lin-29B* specific mutant allele, which may or may not be possible, other experiments could also help to better the understand LIN-29A and LIN-29B activities. Arguably one of the most interesting experiments would be to determine the gene expression changes elicited in animals with different combinations of *lin-29(ΔA); lin-29(ΔAB)* and *mab-10(0)* mutant alleles. To do so, worms of the different genotypes could be harvested in a time course similar to that of the publication in section 2.2, followed by RNA-sequencing. This would reveal which gene expression changes occur in L4 stage animals in all the situations outlined in Figure 12. Therefore, not only gene expression changes specific to LIN-29A, LIN-29B or MAB-10 activity could be extrapolated, but, for instance, also those that are needed for seam cell fusion (comparison of *lin-29(ΔA);mab-10(0)* with *lin-29(ΔAB);mab-10(0)*) or for seam cell exit from the cell cycle (comparison of *lin-29(ΔA);mab-10(0)* with *lin-29(ΔA)*). The latter comparison would also reveal the impact of MAB-10 on LIN-29B activity. This could be very interesting, as MAB-10 may act to enhance or repress transcriptional activity for distinct target genes, as observed for its NAB protein homologs (Sevetson et al., 2000; Svaren et al., 1998). These MAB-10 induced modulations of LIN-29B activity seem to be key for the decision of whether or not to exit the seam cell cycle and could thus reveal interesting gene expression changes that are potentially conserved in mammalian stem cell proliferation or exit from the cell cycle.

3.6 A branched heterochronic pathway regulates the different LIN-29 isoforms

In almost every review or study, a different scheme for the heterochronic pathway is drawn, underlining how unclear the connections between the individual heterochronic genes still are. Major confusion is caused by the observed redundancies between heterochronic genes and by the lack of knowledge about molecular links between the members of the heterochronic pathway (section 1.4). With this work, we can at least resolve part of this confusion, by clarifying the molecular links between *lin-41* and the downstream genes *lin-29* and *mab-10* (section 3.4) and by finding an explanation for the redundancy between *lin-41* and *hbl-1*.

3.6.1 lin-41 and hbl-1 are redundant because they regulate different LIN-29 isoforms

Null mutants of *lin-41* are reported to only display a partial precocious L/A switch at the L3-to-L4 molt (Slack et al., 2000), although both LIN-29A and MAB-10 are precociously expressed (Aeschimann et al., 2017). As we found that both LIN-29 isoforms need to be present for a complete L/A switch to occur (section 2.6), LIN-29B is likely not expressed early enough or to sufficient levels to allow for a complete precocious L/A switch in *lin-41* mutant animals. Interestingly, whereas depletion of only HBL-1 also cannot elicit a full precocious L/A switch, depletion of both LIN41 and HBL-1 together causes a largely complete L/A switch one stage too early (Abrahante et al., 2003; Lin et al., 2003). Therefore, we hypothesized that depletion of HBL-1 results in precocious LIN-29B expression. Indeed, using individually tagged LIN-29 isoforms, we found that depletion of HBL-1 caused a specific upregulation of LIN-29B in seam cell nuclei, whereas depletion of LIN41 led to upregulation of LIN-29A in hyp7 nuclei (section 2.8).

3.6.2 An updated model for the heterochronic pathway

With this result, we cannot only explain the observed redundancy between *lin-41* and *hbl-1*, but also update the model of the heterochronic pathway (Figure 22B,C): Through a yet to be defined mechanism, *hbl-1* specifically represses *lin-29B*, while in a parallel pathway, *lin-41* specifically represses *lin-29A* and *mab-10* through translational repression and mRNA degradation, respectively. The two parallel pathways converge on the execution of the L/A switch, which is only possible when both LIN-29 isoforms and their co-factor MAB-10 are expressed in the epidermis.

Upstream of *lin-41*, the *let-7* miRNA seems to specifically function by repressing *lin-41* (section 3.4.1), whereas the three *let-7* sisters are thought to specifically target *hbl-1* (section 1.4.6). With the worm lines

harboring individually tagged LIN-29 isoforms, we now have a tool to test these proposed specificities in a very clean manner. If *let-7* and the *let-7* sisters really function in two completely different pathways with *lin-41* and *hbl-1* as their respective key targets, we would expect that in *let-7(lf)* mutant animals, only expression of LIN-29A is affected, while in animals mutated in all three *let-7* sisters, only expression of LIN-29B would be changed. In other words, mutations in *let-7* or the three *let-7* sisters should specifically result in a failure of LIN-29A or LIN-29B upregulation in later larval stages, respectively.

3.6.3 Strategies to solve more riddles of the heterochronic pathway

As outlined in section 1.4, a major obstacle to better understand the heterochronic pathway is the lack of knowledge about the molecular interactions between heterochronic genes. With the finding that HBL-1 regulates the expression of LIN-29B, we can propose a potential mechanism. HBL-1 is a transcription factor homologous to fly hunchback and is likely to bind to the promoters of its target genes to regulate transcription. To test a model, in which HBL-1 binds to the promoter of LIN-29B to repress its transcription, Chromatin Immunoprecipitation (ChIP) could be performed, e.g. with a FLAG-tagged HBL-1 protein. However, the observed expression patterns of HBL-1 and LIN-29B would rather support a model in which HBL-1 acts indirectly on LIN-29B expression. After the L1 stage, HBL-1 is only detected in hyp7 and not in seam cells, and its downregulation by the three *let-7* sisters is also specifically observed in hyp7 (Abbott et al., 2005; Abrahante et al., 2003; Lin et al., 2003). By contrast, LIN-29B is specifically upregulated in seam cells upon depletion of HBL-1 (section 2.8). Taken together, this would suggest that *hbl-1* acts indirectly on LIN-29B expression through transmission of a signal from hyp7 to seam cells.

Our finding that the two LIN-29 isoforms are individually regulated by LIN-41 and HBL-1 could be exploited to better define the roles of additional upstream heterochronic genes, such as *lin-4*, *lin-14* and *lin-28*. A particularly confusing and thus interesting heterochronic gene is *lin-28*. As described in section 1.4.7, *lin-28* seems to influence the expression of at least three heterochronic genes, namely *lin-14*, *let-7* and *hbl-1*. Nevertheless, the main phenotype of *lin-28* mutants is a skipping of the L2 stage pattern of seam cell divisions, proposed to be due to its interaction with *hbl-1* (Vadla et al., 2012). By analysing the changes in LIN-29A versus LIN-29B expression in *lin-28* mutants, it could be examined which proposed pathway of LIN-28 is the predominant one at which developmental stage. A major effect due to a failure of suppressing *let-7* maturation should specifically lead to higher LIN-29A levels, while a major effect due to a failure in promoting HBL-1 levels should specifically result in higher LIN-29B levels.

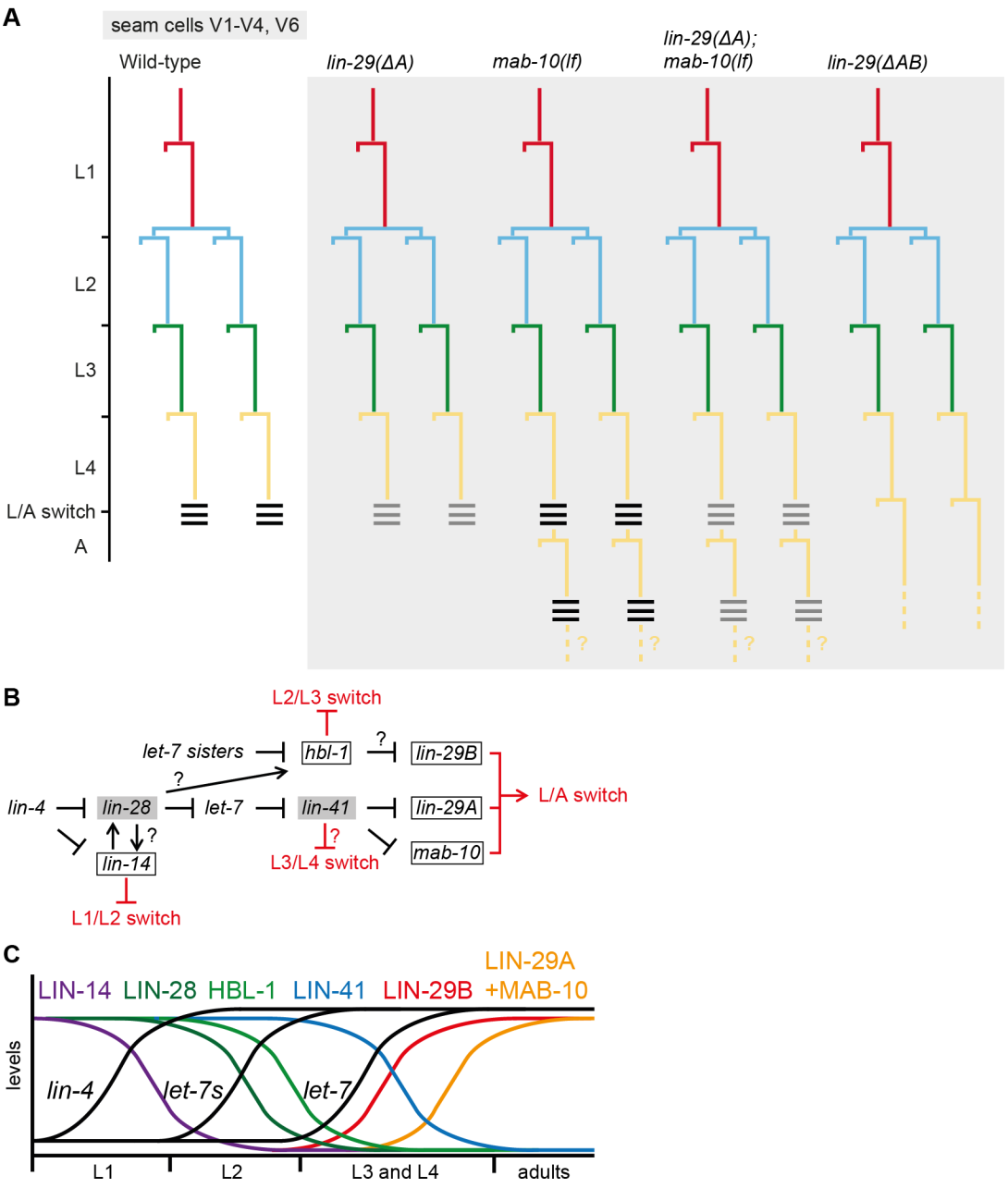


Figure 22. Updated model for the heterochronic pathway

(A) Seam cell lineage patterns for animals with indicated mutations in *lin-29* and/or *mab-10*. Black triple lines indicate wild-type alae structures, grey triple lines indicate weaker alae structures. If seam cell proliferation in *mab-10(lf)* and *lin-29(ΔA); mab-10(lf)* animals ceases after one additional round of extra divisions or not yet remains to be determined (yellow question marks).

(B) Updated model for the heterochronic pathway regulating the switches between the different developmental stages. The proposed regulations of the developmental switches are colored in red, with the link between *lin-41* and the L3/L4 switch being very speculative (red question mark), while unclear molecular links are labelled with a black question mark.

(C) Schematic epidermal gene expression profiles of the heterochronic miRNAs, RBPs and transcription factors or co-factors. The upregulation of mature miRNAs are depicted with black profiles, while the down- or upregulation of transcription factors, transcription co-factors or RBPs are drawn with different colors.

3.7 Extensive conservation of the heterochronic pathway

Several heterochronic genes, as well as their functions in regulating self-renewal and differentiation of stem or progenitor cells, seem to be conserved from worms to mammals (sections 1.3 and 1.5). Most strikingly, mammalian orthologues of *let-7*, *lin-28* and *lin-41* are not only identical to their *C. elegans* counterparts in terms of sequence (*let-7*) or highly similar in terms of domain architecture (*lin-28* and *lin-41*), they are even interconnected in a conserved pathway (section 1.3): LIN28 controls *let-7* processing, and *let-7* regulates LIN41 expression through mRNA degradation and translational repression. The identification of LIN-29A and MAB-10 as the key targets of LIN41 to regulate seam cell self-renewal suggests that the heterochronic pathway is even more extensively conserved than previously thought.

3.7.1 EGR and NAB proteins are conserved regulators of proliferation versus differentiation

LIN-29 contains a domain conserved in Krüppel family early growth response (EGR) transcription factors, and MAB-10 belongs to the family of NAB (NGFI-A-binding protein) transcriptional co-factors that bind to EGR proteins to modulate their activity (Harris and Horvitz, 2011). In mammals, there are four different EGR transcription factors, EGR1-4, and two NAB paralogs, NAB1 and NAB2. Interestingly, EGR and NAB proteins have been described as crucial regulators of proliferation and/or terminal differentiation programs in different animal cell types, including different blood cells or Schwann cells (Du et al., 2014; Laslo et al., 2006; Le et al., 2005; Min et al., 2008; Nguyen et al., 1993; Topilko et al., 1994). For example, macrophage differentiation is promoted by the expression of EGR1, EGR2 and NAB2 (Laslo et al., 2006; Nguyen et al., 1993). In Schwann cells, EGR2, together with NAB1 and NAB2, terminates proliferation of cells in the promyelinating stage and promotes their terminal differentiation to the myelinating stage (Le et al., 2005), processes that are reminiscent of those controlled by LIN-29 and MAB-10 in seam cells. Yet more intriguingly, the function of EGR and NAB proteins in epidermal development could be conserved from worms to mammals, as keratinocytes of mice mutant for both NAB paralogs overproliferate, resulting in a much thicker basal epidermal layer (Le et al., 2005). Finally, EGR1 overexpression was found to decrease the reprogramming efficiency of fibroblasts into iPSCs (Worringer et al., 2014).

3.7.2 Is the regulation of EGRs and NABs through LIN41 conserved?

Since the mammalian counterparts of *lin-28, let-7, lin-41, lin-29* and *mab-10* all regulate self-renewal and differentiation of different stem and progenitor cells, we hypothesize that not only the wiring of the LIN28-*let-7*-LIN41 axis, but also the interconnection between LIN41 and EGRs/NABs could be conserved. In our

proposed model, LIN41 not only inhibits the expression of LIN-29A and MAB-10 in *C. elegans*, but also post-transcriptionally regulates the production of EGR and NAB proteins in mammals. Although direct evidence for a conservation of this pathway to mammals is yet lacking, several observations support this model: First, in the mouse skin, LIN41 is expressed in the basal epidermal layer containing keratinocyte stem cells, and *let-7* is expressed in differentiating keratinocytes of the suprabasal layer, where LIN41 is absent (Rybak et al., 2009). As keratinocytes of mice depleted of NAB1 and NAB2 overproliferate (Le et al., 2005), we hypothesize that LIN41 is downregulated by *let-7* as keratinocytes differentiate, allowing EGR and NAB proteins to be expressed and to stop the proliferation program. Second, *let-7*, LIN41 and EGR1 were all shown to modulate reprogramming efficiency of fibroblasts into iPSCs (Worringer et al., 2014). Strikingly, LIN41 was shown to bind to the EGR1 mRNA in a co-IP experiment in human embryonic stem cells (Worringer et al., 2014). Although a repression mechanism yet remains to be determined, it is possible that LIN41 regulates the translation of EGR1, as EGR1 mRNA levels were only modestly changed upon LIN41 knockdown (Worringer et al., 2014).

In summary, the heterochronic pathway seems to emerge as a fundamental stem cell regulatory pathway in animals. As a part of this pathway, EGR and NAB proteins could be conserved LIN41 targets, with LIN41 regulating their expression at the post-transcriptional level through either translational repression or mRNA degradation (Figure 23). The findings of this thesis could help to provide a mechanistic understanding of how LIN41 achieves regulation of EGR and NAB proteins, enabling targeted experiments in mammalian systems to validate our proposed model.

3.7.3 Strategies to test conservation of LIN41-mediated regulation of EGR/NAB expression

In order to test the model presented in Figure 23, the most important task will be to determine the mechanisms of LIN41-mediated repression on EGR and NAB mRNAs, given that LIN41 indeed regulates EGR and NAB protein production. Some useful tools have been previously established to study LIN41 in mouse embryonic stem (ES) cells (Loedige et al., 2013). These tools include siRNAs for efficient LIN41 knockdown, constructs for expression of a tagged version of LIN41 for co-IPs and luciferase reporter assays to test LIN41-mediated repression on candidate target 3'UTRs. In the long run, it could be very interesting to perform a ribosome profiling time course experiment on mouse ES cells upon knockdown of LIN41, similar to the experiment performed in the publication in section 2.2. This would be an unbiased way of finding targets of LIN41, as determined by identification of the genes earliest upregulated in the time course upon the knockdown. Similar to our experiments in *C. elegans* ((Aeschimann et al., 2017) and section 2.4), the candidate targets could then be verified by LIN41 co-IP experiments. Potential caveats of

these experiments could be a lack of synchrony in the LIN41 knockdown between the different cells, or transfection or knockdown efficiencies in general. Therefore, an alternative strategy would be to perform LIN41 RIP-seq in mouse ES cells. At least in our system in *C. elegans*, RIP-seq revealed the same set of LIN41 targets as the ribosome profiling time course experiments ((Aeschimann et al., 2017) and section 2.4), suggesting that this strategy could prove successful.

In the short run, more specific and targeted experiments could be performed to test if LIN41 represses EGR and NAB mRNAs: The 3'UTRs and 5'UTRs of the different EGR and NAB mRNAs could be cloned into a reporter system, for example the already established luciferase system (Loedige et al., 2013). After transfection into mouse ES cells, the reporters could be tested for upregulation upon LIN41 knockdown. Furthermore, by measuring both luciferase activity and mRNA levels of the reporters, it could be determined if repression occurs on the translational level or through mRNA degradation. EGR2 seems to be an interesting candidate for being translationally repressed through its 5'UTR, as some of its annotated 5'UTR variants consist of an upstream 5'UTR exon spliced to a downstream exon with the ORF start codon, reminiscent of the *lin-29A* mRNA structure.

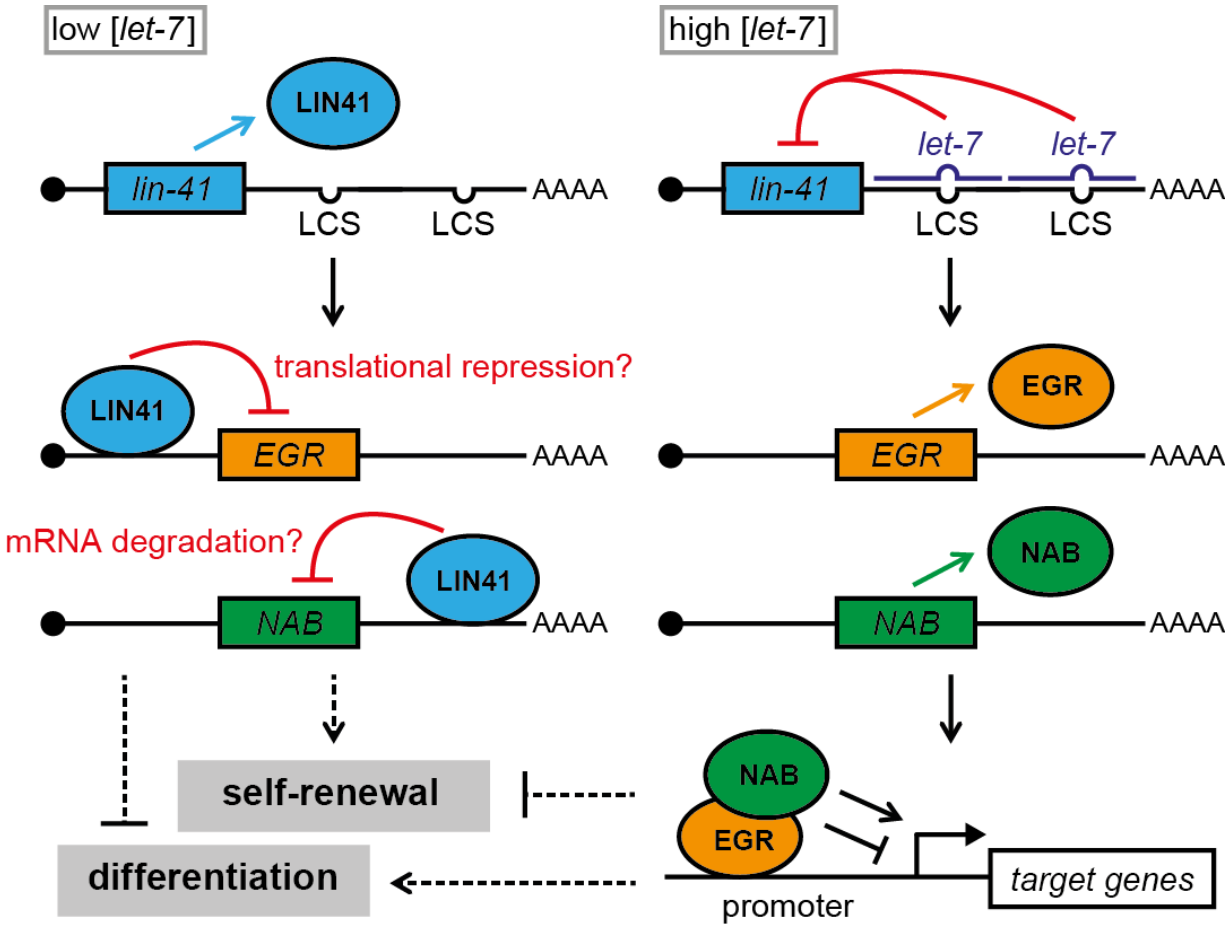


Figure 23. The *let-7*-LIN41 pathway as a conserved stem cell regulatory pathway.
Model for a potentially conserved pathway to control self-renewal and differentiation of mammalian stem cells. Whereas *let-7* and LIN41, as well as EGR and NAB proteins, have been shown to regulate stem cell development, the interconnections between these factors are less clear. On the one hand, it is known that *let-7* regulates LIN41, in *C. elegans* as well as in mammals. On the other hand, future experiments will have to determine whether LIN41 indeed regulates EGR and NAB protein expression, and if so, through which mechanisms.

4 Supplemental Methods

This section describes methods and worm strains used in this thesis but not described in the publication in section 2.2.

Seam cell imaging and quantification.

Arrested L1 larvae were plated on 2% NGM agar plates with *Escherichia coli* OP50 bacteria (Stiernagle, 2006). Synchronized worms were grown for 36-38 hours (late L4 stage) or 40-42 hours (young adult stage), with the exact developmental time assessed by staging of individual worms according to gonad length and vulva morphology. Worms were mounted to a 2 % (w/v) agarose pad and immobilized in 10 mM levamisole. Fluorescent and Differential Interference Contrast (DIC) images were acquired with a Zeiss Axio Observer Z1 microscope using the AxioVision SE64 software. Selections of regions and processing of images was performed with Fiji (Schindelin et al., 2012). Seam cell quantifications were performed by counting all clearly visible fluorescent cells (*scm::gfp* transgene) of the upper lateral side in mounted worms.

Generation of a balancer for *lin-41(xe8)* using CRISPR-Cas9.

The *lin-41(xe8)* allele is not temperature-sensitive like *let-7(n2853)* and therefore causes lethality at any temperature (Ecsedi et al., 2015). In order to maintain *lin-41(xe8)* animals, a balancer null allele, *lin-41(bch28)*, was previously created by inserting an expression cassette driving ubiquitous nuclear GFP from the *eft-3* promoter into the *lin-41* coding sequence (Katic et al., 2015). To avoid recombination of *lin-41(bch28)* with *lin-41(xe8)*, generating a wild-type *lin-41* copy and a recombined *lin-41(bch28 xe8)* allele, we additionally deleted a large part of the *lin-41* coding sequence together with the part of the *lin-41* 3'UTR containing the *let-7* complementary sites within the *lin-41(bch28)* allele. To this end, *lin-41(bch28)* heterozygous worms were injected with a mix containing 50 ng/µl pIK155, 100 ng/µl of each pIK198 with a cloned sgRNA, 5 ng/µl pCFJ90 and 5 ng/µl pCFJ104, as previously described (Katic et al., 2015). We injected two sgRNAs, sgRNA1 (ggtgactgaatcattgacgg) and sgRNA2 (agaaggtttcaatggttcag), cutting in the third coding exon and the 3'UTR of *lin-41*, respectively. Single F1 progeny of injected wild-type worms were picked to individual plates and the F2 progeny were screened for expected deletions in *lin-41(bch28)* by PCR. The obtained allele *lin-41(bch28 xe70)* was further confirmed by DNA sequencing and outcrossed three times to the wild-type strain before crossing it with *lin-41(xe8)* heterozygous animals. The final *lin-41(bch28 xe70)* allele consists of the inserted expression cassette, as described in (Katic et al., 2015), followed by an additional deletion of the region with the following flanking sequences: 5' ggctcactatttgacactcc – *xe70* deletion (6395 bp) – accattgaaaccttctccc 3'.

Generation of novel *lin-29*, *lin-29A* and *mab-10* null mutant alleles using CRISPR-Cas9

Wild-type worms were injected with a mix containing 50 ng/µl pIK155, 100 ng/µl of each pIK198 with a cloned sgRNA, 5 ng/µl pCFJ90 and 5 ng/µl pCFJ104, as previously described (Katic et al., 2015). Single F1 progeny of injected wild-type worms were picked to individual plates and the F2 progeny screened for deletions using PCR assays. After analysis by DNA sequencing, the alleles were outcrossed three times to the wild-type strain.

In order to obtain mutant alleles for *lin-29* that were undoubtedly null, we generated large deletions lacking almost the entire coding region and chose two alleles, *xe36* and *xe37*, for backcrossing and characterization. The two alleles were obtained by injecting two sgRNAs, sgRNA1 (gctggaaccaccactggctc) and sgRNA2 (atattatttatcagtgattg), cutting in the first and last coding exons of *lin-29A*, respectively. The resulting *lin-29(xe36)* allele is a 14,783 bp deletion with

a 3 bp insertion with the following flanking sequences: 5' ggaatagctggaaccaccac – *xe36* deletion – *xe36* insertion (gga) – cactgataaataatatgaaa 3'. The resulting *lin-29(xe37)* allele is a 14,801 bp deletion with a 2 bp insertion with the following flanking sequences: 5' ggactctggaatagctggaa – *xe37* deletion – *xe37* insertion (aa) – aatatgaaaaatcattccta 3'. Translation of *xe36* and *xe37* yield only short stretches of 27 and 28 amino acids (xe36: MDQTVLDSAFNSPVDSGIAGTTT-DTDK*, xe37: MDQTVLDSAFNSPVDSGIAG-KNMKNHSY*), containing the N-terminal 23 and 20 amino acids of LIN-29A, respectively. The small insertions then lead to translation of an additional amino acid (D and K, respectively), followed by a short stretch 3 out-of-frame and 8 in-frame amino acids, respectively.

In order to specifically mutate *lin-29A* without affecting expression of *lin-29B*, we introduced deletions into the coding exons specific to *lin-29A*, to introduce a frame-shift in the downstream *lin-29A* reading frame without affecting the lin-29B open reading frame. Of the many *lin-29A* mutants obtained by Cas9-CRISPR, four alleles were backcrossed and characterized, as described below. For most experiments, the *lin-29A(xe40)* allele was used. This 1102 bp deletion removes a large part of the coding exons specific to *lin-29A*, at the same time introducing a frame-shift in the *lin-29A* reading frame with a predicted stop codon in exon 6. The deletion has the following flanking sequences: 5' ctctggaatagctggaaccac – *xe40* deletion – attctctcctgccacatcat 3'. Translation of *xe40* yields a protein with the N-terminal 22 amino acids of LIN-29A (MDQTVLDSAFNSPVDSGIAGTT), followed by a stretch of 69 out-of-frame amino acids and a stop codon. *xe40* was obtained by injecting two sgRNAs, one (gctggaaccaccactggctc) cutting in exon 2 right downstream of the *lin-29A* ATG start codon, the other one (gtggcaggagagaattctga) cutting in exon 4, the most downstream *lin-29A* specific exon. With the same sgRNAs as used to get *xe40*, two other alleles, *xe39* and *xe42*, were obtained. The resulting *lin-29A(xe39)* allele is a 1095 bp deletion with the following flanking sequences: 5' atagctggaaccaccactgg – *xe39* deletion – aattctctcctgccacatca 3'. Translation of *xe39* yields a protein with the N-terminal 24 amino acids of LIN-29A (MDQTVLDSAFNSPVDSGIAGTTTG), followed by a stretch of 57 out-of-frame amino acids and a stop codon. The resulting *lin-29A(xe42)* allele is a 1077 bp deletion with a 19 bp insertion with the following flanking sequences: 5' gaatagctggaaccaccact – *xe42* deletion – *xe42* insertion (accacccattagatcaccc) – atctactgaacttccatcag 3'. Translation of *xe42* yields a protein with the N-terminal 23 amino acids of LIN-29A (MDQTVLDSAFNSPVDSGIAGTTT), followed by a stretch of 3 out-of-frame amino acids and a stop codon within the insertion (underlined). A smaller deletion, *xe38*, was obtained through injection of only the sgRNA targeting exon 2 (gctggaaccaccactggctc). The *lin-29A(xe38)* allele is a 10 bp deletion with the following flanking sequences: 5' gaatagctggaaccaccact – *xe38* deletion – ctaccacccattttggtgtt 3'. Translation of *xe38* yields a protein with the N-terminal 23 amino acids of LIN-29A (MDQTVLDSAFNSPVDSGIAGTTT), followed by a stretch of 14 out-of-frame amino acids and a stop codon in exon 3.

In order to obtain a mutant allele for *mab-10* that was undoubtedly null, we generated a large deletion (*xe44*) lacking almost the entire coding region. The *mab-10(xe44)* allele was obtained by injecting two sgRNAs, sgRNA1 (gatgatgatgatgaagaggt) and sgRNA2 (gctcccggaatcttgaagct), cutting in the second coding exon and the beginning of the 3'UTR of *mab-10*, respectively. The resulting *xe44* allele is a 2901 bp deletion with a 4 bp insertion with the following flanking sequences: 5' ttatcatctcttacaactca – *xe44* deletion – *xe44* insertion (ctct) – tatttttttgttttcctcgtga 3'. Translation of *xe44* yields a 58 amino acid stretch (MSSSSSSSLPTSSASTTTSSITSRPSASHHLESILSSSSSSPSILSSLTT-HSYFLFSS*) containing the N-terminal 50 amino acids of MAB-10, followed by 8 additional amino acids, translated from the small insertion and the *mab-10* 3'UTR, and a stop codon (underlined in the flanking sequence above). When injecting the sgRNAs to obtain the *xe44* allele, a second *mab-10(null)* allele, *xe43*, was backcrossed and characterized. The *xe43* allele is a 2886 bp deletion with the following flanking sequences: 5' ttacaactcaatccacttca – *xe43* deletion – ttccatatttttttgttttcc 3'. Translation of *xe43* yields in a 77 amino acid stretch (MSSSSSSSLPTSSASTTTSSITSRPSASHHLESILSSSSSSPSILSSLTTQSTSFHIFCFPRDLCNTS-RWSVSCDSH*) containing the N-terminal 54 amino acids of MAB-10, followed by 23 additional amino acids, translated from the *mab-10* 3'UTR, and a stop codon.

Isoform-specific tagging of endogenous *lin-29* using CRISPR-Cas9.

In order to specifically tag *lin-29A* at the N-terminus, the following mix was injected into wild-type worms (Dickinson et al., 2015; Katic et al., 2015): 50 ng/µl pIK155, 100 ng/µl of pIK198 with a cloned sgRNA (atattatttatcagtgattg), 2.5 ng/µl pCFJ90, 5 ng/µl pCFJ104 and 10 ng/µl pDD282 with cloned homology arms as described below. For specific tagging of *lin-29B* at the C-terminus, the following mix was injected into *lin-29A(xe40)* mutant worms: 50 ng/µl pIK155, 100 ng/µl of pIK198 with a cloned sgRNA (atattatttatcagtgattg), 2.5 ng/µl pCFJ90, 5 ng/µl pCFJ104 and 10 ng/µl pDD282 with cloned homology arms as described in the publication in section 2.2. For N-terminal tagging of *lin-29A*, the plasmid for homologous recombination was prepared by restriciton digest of pDD282 with ClaI and SpeI, followed by a Gibson assembly reaction (Gibson et al., 2009) with two gBlocks® Gene Fragments (Integrated DNA Technologies):

gBlock1:

AGTCACGACGTTGTAAAACGACGGCCAGTCGCCGGCAgtgaattcgaacgtttcgtggcaggtggtgggcacgtgtaatcggaaccgagtaaaagg
gaatcgaaaggagtagacaaagggcatataagaaatctttcagttcttctttgggtcacatggttttaggagatttttttcttgtcttgaatttcattagattaaagtattc
gttaggtacaataatgctacactttccaagtttcaccgccgtacttcggggcaggtgcagtatgttgttcctcgacttgtgaaatccaagtctatgcaaaaacaaaaaa
aaatcaacatagaccacctgccgtgttttccgattacggtatgagtatgagttttaggccccacctctgttaattaacttttttcctatagtatttttttttctcaaagtgtag
attattagcaatgtggccgtgtaatttttatgggtcttcccaaaatttttggagccaggtaatatttaattttacagccgatgggttactgtattaactaggaaatggtaga
aaaagattttacgaatattttgttagttgcctatttttttagtcgaagaaaaccacaactgagatataagctataatcaaagttgagaataaaaattaatccaagtagat
gttttaggcatttaatctataaagaaatatatactttcctttcagcttcgcgaggatcaagccaacttcttcaacgcaATGAGTAAAGGAGAAGAATTGTTC
ACTGGAGTTGTCCCA.

gBlock2:

CGACGACAAGCGTGATTACAAGGATGACGATGACAAGAGAatggatcaaactgttctagattcggcattcaactctccagtggactctggaatagc
tggaaccac**G**ac**A**ggc**AGT**ggatctaccacccattttggtgttggaacaaactgtgaggatttgattatttcaagaaactgatcataaaaaataaagttaaagtttc
agtgagatcatcaagtcggtcaaccgacggcactgactcaacggacggtgccaactcggacaatgtgaccggatcaactggatccacaccagcacatcattcgatca
ctaatttgaatatggctttgtcgcaacactccattgactcggcaactgctgcttcatcgacaaacccgtttccgcattttaatcaaggtagatttatcaaagcttttgtttc
gaaaatttcatttaaattcaatttctgtatcttcaaagaatagaaaatgcgatcaattttgatagcagaaaacttaaaactgagattgagcctgcgcctatgcttaaatc
aaagaactgaactttagacaaagcctaagaaaaatcccatgtctaggtccaaatccaagcctgaacctgagcccaggtcttccgaaggtcgtcggtaagatattattt
tagactctagtagccaattacttcatctacccgtacCGATAACATGGTCATAGCTGTTTCCTGTGTGAAATTG. The overhangs for Gibson assembly are written in capital letters, regions homologous to the *lin-29* locus in lowercase letters and introduced silent mutations in bold underlined capital letters. Recombinants were isolated according to the protocol by Dickinson *et al*. (Dickinson et al., 2015), verified by DNA sequencing and outcrossed three times. From each injection, two independent worm lines were obtained and characterized.

Additional worm strains used in this thesis

Strain number or reference	Genotype
HW4	*lin-29(n546)* II
HW1407	*lin-29(n836)* II
HW1691	*lin-29(xe36)* II
HW1692	*lin-29(xe37)* II
HW1693	*lin-29a(xe38)* II
HW1694	*lin-29a(xe39)* II
HW1695	*lin-29a(xe40)* II
HW1696	*lin-29a(xe42)* II
HW1697	*mab-10(xe43)* II
HW1698	*mab-10(xe44)* II
HW1789	*lin-29(xe36) mab-10(xe43)* / mnC1 II
HW1790	*lin-29(xe37) mab-10(xe44)* / mnC1 II
HW1755	*lin-29(xe38) mab-10(xe43)* / mnC1 II
HW1756	*lin-29(xe40) mab-10(xe43)* / mnC1 II
HW1757	*lin-29(xe38) mab-10(xe44)* / mnC1 II
HW1758	*lin-29(xe40) mab-10(xe44)* / mnC1 II
HW1665	*lin-41(bch28[Peft-3::gfp::h2b::tbb-2 3'UTR])/lin-41(xe8)* I
HW1870	*lin-41(bch28[Peft-3::gfp::h2b::tbb-2 3'UTR] xe70)/lin-41(xe8)* I
HW647	*wIs54[scm::gfp]* V; *mjIs15[ajm-1::mCherry]*

HW1387	*wIs54[scm::gfp]* V; *mjIs15[ajm-1::mCherry]*; *let-7(n2853)* X
HW1865	*lin-41(xe8) / lin-41(bch28 xe70)* I; *wIs54[scm::gfp]* V; *mjIs15[ajm-1::mCherry]*
HW1409	*lin-41(xe11)* I; *wIs54[scm::gfp]* V; *let-7(n2853)* X
HW1861	*lin-29a(xe40)* II; *wIs54[scm::gfp]* V; *mjIs15[ajm-1::mCherry]*
HW1862	*mab-10(xe44)* II; *wIs54[scm::gfp]* V; *mjIs15[ajm-1::mCherry]*
HW1864	*lin-29a(xe40) mab-10(xe44)* / mnC1 II; *wIs54[scm::gfp]* V; *mjIs15[ajm-1::mCherry]*
HW1860	*lin-29(xe37)* / mnC1 II; *wIs54[scm::gfp]* V; *mjIs15[ajm-1::mCherry]*
HW1863	*lin-29(xe37) mab-10(xe44)* / mnC1 II; *wIs54[scm::gfp]* V; *mjIs15[ajm-1::mCherry]*
HW1826	*lin-29(xe63[gfp::3xflag::lin-29A])* II
HW1834	*lin-29(xe64[gfp::3xflag::lin-29A])* II (alternative independent CRISPR integrant)
HW1835	*lin-29(xe65[lin-29A(xe40)::gfp::3xflag])* II
HW1842	*lin-29(xe66[lin-29A(xe40)::gfp::3xflag])* II (alternative independent CRISPR integrant)
HW1508	*him-5(e1490)* V
HW1814	*lin-41(xe8) / lin-41(bch28)* I*; him-5(e1490)* V
HW1616	*lin-41(xe11)* I; *him-5(e1490)*
HW1617	*lin-41(xe11)* I; *him-5(e1490)* V; *let-7(n2853)* X
HW1618	*him-5(e1490)* V; *let-7(n2853)* X

5 Supplemental Tables

Tables S1-S4: Quantification of vulva phenotypes

Quantification of burst, Pvl and Egl worms grown at 25 °C (Tables S1-S3) or burst worms grown at 15 °C (Table S4). Three biologically independent experiments were performed and at least 400 worms counted for each condition. Indicated is the percentage of counted worms with the respective phenotype.

Table S1

Genotype	**% burst worms experiment 1**	**% burst worms experiment 2**	**% burst worms experiment 3**	**% burst worms average**
wild-type	0.0	0.0	0.0	0.0
let-7(n2853)	92.0	90.7	85.3	89.3
lin-41(xe8)	86.9	89.9	88.8	88.5
mab-10(xe44)	0.2	0.0	0.2	0.2
lin-29(xe40)	1.7	1.6	2.1	1.8
lin-29(xe37)	2.3	1.0	2.1	1.8
lin-29(xe37); mab-10(xe44)	1.9	1.2	1.6	1.6
lin-29(xe40); mab-10(xe44)	27.5	29.1	29.5	28.7

Table S2

Genotype	**% Pvl worms experiment 1**	**% Pvl worms experiment 2**	**% Pvl worms experiment 3**	**% Pvl worms average**
wild-type	0.0	0.0	0.0	0.0
mab-10(xe44)	0.2	0.0	0.2	0.2
lin-29(xe40)	68.5	64.2	53.3	62.0
lin-29(xe37)	98.6	99.6	99.1	99.1
lin-29(xe37); mab-10(xe44)	96.8	98.3	98.6	97.9
lin-29(xe40); mab-10(xe44)	96.9	98.5	97.8	97.7

Table S3

Genotype	**% Egl worms experiment 1**	**% Egl worms experiment 2**	**% Egl worms experiment 3**	**% Egl worms average**
wild-type	0.0	0.0	0.0	0.0
mab-10(xe44)	6.6	4.2	5.9	5.6
lin-29(xe40)	82.6	83.9	79.7	82.1
lin-29(xe37)	99.0	99.3	99.6	99.3
lin-29(xe37); mab-10(xe44)	97.3	98.6	97.7	97.9
lin-29(xe40); mab-10(xe44)	97.1	98.9	98.3	98.1

Table S4

Genotype	% burst worms experiment 1	% burst worms experiment 2	% burst worms experiment 3	% burst worms average
wild-type	0.0	0.0	0.0	0.0
let-7(n2853)	3.9	5.7	11.9	7.2
lin-41(xe8)	86.5	90.1	94.5	90.4
mab-10(xe44)	0.0	0.2	0.0	0.1
lin-29(xe40)	1.9	2.5	4.3	2.9
lin-29(xe37)	6.8	17.9	22.5	15.7
lin-29(xe37); mab-10(xe44)	5.4	16.0	11.7	11.0
lin-29(xe40); mab-10(xe44)	50.4	46.4	52.1	49.6

Tables S5-S7: Quantification of seam cell numbers in late L4 stage worms

Quantification of seam cell numbers in worms of indicated genotypes, at the late L4 stage (Table S5), the young adult stage (Table S6) and the later adult stage (Table S7). Indicated is the percentage of counted worms with the respective number of seam cells.

Table S5

worm line with SCM::GFP	Number of seam cells																		Worms counted
	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	
wild-type	5	60	35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
let-7(n2853)	0	65	20	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
lin-41(xe8)	0	80	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
mab-10(xe44)	0	65	30	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
lin-29(xe40)	0	65	35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
lin-29(xe37)	0	80	15	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
lin-29(xe37); mab-10(xe44)	0	85	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
lin-29(xe40); mab-10(xe44)	0	65	20	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
lin-41(xe11); let-7(n2853)	0	55	30	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20

Table S6

worm line with SCM::GFP	Number of seam cells																		Worms counted
	15	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	
wild-type	0	69	25	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	52
let-7(n2853)	2	17	15	4	0	0	4	2	0	0	10	13	15	13	0	2	0	2	52
lin-41(xe8)	2	20	13	0	0	0	2	2	4	0	11	17	9	7	11	0	2	0	54
mab-10(xe44)	2	57	34	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53
lin-29(xe40)	2	65	27	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	52
lin-29(xe37)	0	0	0	0	0	0	0	0	0	0	2	18	24	25	12	10	8	2	51
lin-29(xe37); mab-10(xe44)	0	0	0	0	0	0	0	0	0	0	6	12	23	25	17	6	4	8	52
lin-29(xe40); mab-10(xe44)	0	23	16	4	0	0	0	2	2	0	13	13	9	11	2	4	2	2	56
lin-41(xe11); let-7(n2853)	2	54	31	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	52

Table S7

worm line with SCM::GFP	Number of seam cells																		Worms counted
	15	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	
wild-type	0	50	32	14	5	0	0	0	0	0	0	0	0	0	0	0	0	0	22
mab-10(xe44)	4	21	25	4	4	17	4	17	4	0	0	0	0	0	0	0	0	0	24
lin-29(xe40)	0	50	40	5	5	0	0	0	0	0	0	0	0	0	0	0	0	0	20
lin-29(xe37)	0	15	15	0	0	10	5	0	0	15	10	10	5	10	5	0	0	0	20
lin-29(xe37); mab-10(xe44)	0	20	5	0	0	10	10	5	0	10	20	10	0	10	0	0	0	0	20
lin-29(xe40); mab-10(xe44)	0	0	0	0	0	0	0	0	0	5	5	23	14	27	14	9	0	5	22
lin-41(xe11); let-7(n2853)	0	60	25	10	5	0	0	0	0	0	0	0	0	0	0	0	0	0	20

Tables S8-S9: Quantification of alae structures

Quantification of alae structures in worms of indicated genotypes, at the young adult stage (Table S8) and the later adult stage (Table S9). Indicated is the percentage of counted worms with the respective alae structure.

Table S8

Genotype	Complete wt alae	Complete weak alae	Partial weak alae ≥ 50%	Partial weak alae < 50%	Absent alae	Worms counted
wild-type	100	0	0	0	0	32
let-7(n2853)	0	9.4	12.5	12.5	65.6	32
lin-41(xe8)	0	9.4	15.6	12.5	62.5	32
mab-10(xe44)	100	0	0	0	0	32
lin-29(xe40)	0	63.6	36.4	0	0	33
lin-29(xe37)	0	0	0	0	100	32
lin-29(xe37); mab-10(xe44)	0	0	0	0	100	32
lin-29(xe40); mab-10(xe44)	0	55.9	44.1	0	0	34

Table S9

Genotype	Complete wt alae	Complete weak alae	Partial weak alae ≥ 50%	Partial weak alae < 50%	Absent alae	Worms counted
wild-type	100	0	0	0	0	20
mab-10(xe44)	100	0	0	0	0	20
lin-29(xe40)	0	100	0	0	0	20
lin-29(xe37)	0	0	0	0	100	20
lin-29(xe37); mab-10(xe44)	0	0	0	0	100	20
lin-29(xe40); mab-10(xe44)	0	100	0	0	0	20

6 References

Abaza, I., Coll, O., Patalano, S., and Gebauer, F. (2006). Drosophila UNR is required for translational repression of male-specific lethal 2 mRNA during regulation of X-chromosome dosage compensation. Genes Dev *20*, 380-389.

Abbott, A.L., Alvarez-Saavedra, E., Miska, E.A., Lau, N.C., Bartel, D.P., Horvitz, H.R., and Ambros, V. (2005). The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in Caenorhabditis elegans. Dev Cell *9*, 403-414.

Abrahante, J.E., Daul, A.L., Li, M., Volk, M.L., Tennessen, J.M., Miller, E.A., and Rougvie, A.E. (2003). The Caenorhabditis elegans hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs. Dev Cell *4*, 625-637.

Aeschimann, F., Kumari, P., Bartake, H., Gaidatzis, D., Xu, L., Ciosk, R., and Grosshans, H. (2017). LIN41 Post-transcriptionally Silences mRNAs by Two Distinct and Position-Dependent Mechanisms. Mol Cell *65*, 476-489 e474.

Aeschimann, F., Xiong, J., Arnold, A., Dieterich, C., and Grosshans, H. (2015). Transcriptome-wide measurement of ribosomal occupancy by ribosome profiling. Methods *85*, 75-89.

Ambros, V. (1989). A hierarchy of regulatory genes controls a larva-to-adult developmental switch in C. elegans. Cell *57*, 49-57.

Ambros, V., and Horvitz, H.R. (1984). Heterochronic mutants of the nematode Caenorhabditis elegans. Science *226*, 409-416.

Ambros, V., and Horvitz, H.R. (1987). The lin-14 locus of Caenorhabditis elegans controls the time of expression of specific postembryonic developmental events. Genes Dev *1*, 398-414.

Arasu, P., Wightman, B., and Ruvkun, G. (1991). Temporal regulation of lin-14 by the antagonistic action of two other heterochronic genes, lin-4 and lin-28. Genes Dev *5*, 1825-1833.

Bag, J. (2001). Feedback inhibition of poly(A)-binding protein mRNA translation. A possible mechanism of translation arrest by stalled 40 S ribosomal subunits. J Biol Chem *276*, 47352-47360.

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol *2*, 28-36.

Baltz, A.G., Munschauer, M., Schwanhausser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M.*, et al.* (2012). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell *46*, 674-690.

Balzer, E., Heine, C., Jiang, Q., Lee, V.M., and Moss, E.G. (2010). LIN28 alters cell fate succession and acts independently of the let-7 microRNA during neurogliogenesis in vitro. Development *137*, 891-900.

Balzer, E., and Moss, E.G. (2007). Localization of the developmental timing regulator Lin28 to mRNP complexes, P-bodies and stress granules. RNA Biol *4*, 16-25.

Barreau, C., Paillard, L., and Osborne, H.B. (2005). AU-rich elements and associated factors: are there unifying principles? Nucleic Acids Res *33*, 7138-7150.

Bashaw, G.J., and Baker, B.S. (1997). The regulation of the Drosophila msl-2 gene reveals a function for Sex-lethal in translational control. Cell *89*, 789-798.

Bazzini, A.A., Lee, M.T., and Giraldez, A.J. (2012). Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. Science *336*, 233-237.

Beckmann, K., Grskovic, M., Gebauer, F., and Hentze, M.W. (2005). A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in Drosophila. Cell *122*, 529-540.

Bethune, J., Artus-Revel, C.G., and Filipowicz, W. (2012). Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells. EMBO Rep *13*, 716-723.

Bettinger, J.C., Euling, S., and Rougvie, A.E. (1997). The terminal differentiation factor LIN-29 is required for proper vulval morphogenesis and egg laying in Caenorhabditis elegans. Development *124*, 4333-4342.

Bettinger, J.C., Lee, K., and Rougvie, A.E. (1996). Stage-specific accumulation of the terminal differentiation factor LIN-29 during Caenorhabditis elegans development. Development *122*, 2517-2527.

Brooks, S.A., and Blackshear, P.J. (2013). Tristetraprolin (TTP): interactions with mRNA and proteins, and current thoughts on mechanisms of action. Biochim Biophys Acta *1829*, 666-679.

Burkhard, P., Stetefeld, J., and Strelkov, S.V. (2001). Coiled coils: a highly versatile protein folding motif. Trends Cell Biol *11*, 82-88.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M.*, et al.* (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell *149*, 1393-1406.

Chang, H., Lim, J., Ha, M., and Kim, V.N. (2014). TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. Mol Cell *53*, 1044-1052.

Chang, H.M., Martinez, N.J., Thornton, J.E., Hagan, J.P., Nguyen, K.D., and Gregory, R.I. (2012). Trim71 cooperates with microRNAs to repress Cdkn1a expression and promote embryonic stem cell proliferation. Nat Commun *3*, 923.

Chekulaeva, M., Mathys, H., Zipprich, J.T., Attig, J., Colic, M., Parker, R., and Filipowicz, W. (2011). miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. Nat Struct Mol Biol *18*, 1218-1226.

Chen, J., Lai, F., and Niswander, L. (2012). The ubiquitin ligase mLin41 temporally promotes neural progenitor cell maintenance through FGF signaling. Genes Dev *26*, 803-815.

Chen, Y., Boland, A., Kuzuoglu-Ozturk, D., Bawankar, P., Loh, B., Chang, C.T., Weichenrieder, O., and Izaurralde, E. (2014). A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. Mol Cell *54*, 737-750.
Chen, Y.L., Yuan, R.H., Yang, W.C., Hsu, H.C., and Jeng, Y.M. (2013). The stem cell E3-ligase Lin-41 promotes liver cancer progression through inhibition of microRNA-mediated gene silencing. J Pathol *229*, 486-496.
Cho, J., Chang, H., Kwon, S.C., Kim, B., Kim, Y., Choe, J., Ha, M., Kim, Y.K., and Kim, V.N. (2012). LIN28A is a suppressor of ER-associated translation in embryonic stem cells. Cell *151*, 765-777.
Cho, P.F., Gamberi, C., Cho-Park, Y.A., Cho-Park, I.B., Lasko, P., and Sonenberg, N. (2006). Cap-dependent translational inhibition establishes two opposing morphogen gradients in Drosophila embryos. Curr Biol *16*, 2035-2041.
Cho, P.F., Poulin, F., Cho-Park, Y.A., Cho-Park, I.B., Chicoine, J.D., Lasko, P., and Sonenberg, N. (2005). A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. Cell *121*, 411-423.
Chou, C.F., Mulky, A., Maitra, S., Lin, W.J., Gherzi, R., Kappes, J., and Chen, C.Y. (2006). Tethering KSRP, a decay-promoting AU-rich element-binding protein, to mRNAs elicits mRNA decay. Mol Cell Biol *26*, 3695-3706.
Cianetti, L., Segnalini, P., Calzolari, A., Morsilli, O., Felicetti, F., Ramoni, C., Gabbianelli, M., Testa, U., and Sposi, N.M. (2005). Expression of alternative transcripts of ferroportin-1 during human erythroid differentiation. Haematologica *90*, 1595-1606.
Conrad, T., and Akhtar, A. (2011). Dosage compensation in Drosophila melanogaster: epigenetic fine-tuning of chromosome-wide transcription. Nat Rev Genet *13*, 123-134.
Cuevas, E., Rybak-Wolf, A., Rohde, A.M., Nguyen, D.T., and Wulczyn, F.G. (2015). Lin41/Trim71 is essential for mouse development and specifically expressed in postnatal ependymal cells of the brain. Front Cell Dev Biol *3*, 20.
Deiuliis, J.A. (2016). MicroRNAs as regulators of metabolic disease: pathophysiologic significance and emerging role as biomarkers and therapeutics. Int J Obes (Lond) *40*, 88-101.
Del Rio-Albrechtsen, T., Kiontke, K., Chiou, S.Y., and Fitch, D.H. (2006). Novel gain-of-function alleles demonstrate a role for the heterochronic gene lin-41 in C. elegans male tail tip morphogenesis. Dev Biol *297*, 74-86.
Dickinson, D.J., Pani, A.M., Heppert, J.K., Higgins, C.D., and Goldstein, B. (2015). Streamlined Genome Engineering with a Self-Excising Drug Selection Cassette. Genetics *200*, 1035-1049.
Ding, X.C., and Grosshans, H. (2009). Repression of C. elegans microRNA targets at the initiation level of translation requires GW182 proteins. EMBO J *28*, 213-222.
Djuranovic, S., Nahvi, A., and Green, R. (2012). miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. Science *336*, 237-240.
Du, N., Kwon, H., Li, P., West, E.E., Oh, J., Liao, W., Yu, Z., Ren, M., and Leonard, W.J. (2014). EGR2 is critical for peripheral naive T-cell differentiation and the T-cell response to influenza. Proc Natl Acad Sci U S A *111*, 16484-16489.
Duncan, K., Grskovic, M., Strein, C., Beckmann, K., Niggeweg, R., Abaza, I., Gebauer, F., Wilm, M., and Hentze, M.W. (2006). Sex-lethal imparts a sex-specific function to UNR by recruiting it to the msl-2 mRNA 3' UTR: translational repression for dosage compensation. Genes Dev *20*, 368-379.
Ecsedi, M., and Grosshans, H. (2013). LIN-41/TRIM71: emancipation of a miRNA target. Genes Dev *27*, 581-589.
Ecsedi, M., Rausch, M., and Grosshans, H. (2015). The let-7 microRNA directs vulval development through a single target. Dev Cell *32*, 335-344.
Eichhorn, S.W., Guo, H., McGeary, S.E., Rodriguez-Mias, R.A., Shin, C., Baek, D., Hsu, S.H., Ghoshal, K., Villen, J., and Bartel, D.P. (2014). mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues. Mol Cell *56*, 104-115.
Euling, S., Bettinger, J.C., and Rougvie, A.E. (1999). The LIN-29 transcription factor is required for proper morphogenesis of the Caenorhabditis elegans male tail. Dev Biol *206*, 142-156.
Fagoonee, S., Bearzi, C., Di Cunto, F., Clohessy, J.G., Rizzi, R., Reschke, M., Tolosano, E., Provero, P., Pandolfi, P.P., Silengo, L*., et al.* (2013). The RNA binding protein ESRP1 fine-tunes the expression of pluripotency-related factors in mouse embryonic stem cells. PLoS One *8*, e72300.
Friend, K., Campbell, Z.T., Cooke, A., Kroll-Conner, P., Wickens, M.P., and Kimble, J. (2012). A conserved PUF-Ago-eEF1A complex attenuates translation elongation. Nat Struct Mol Biol *19*, 176-183.
Fukao, A., Mishima, Y., Takizawa, N., Oka, S., Imataka, H., Pelletier, J., Sonenberg, N., Thoma, C., and Fujiwara, T. (2014). MicroRNAs trigger dissociation of eIF4AI and eIF4AII from target mRNAs in humans. Mol Cell *56*, 79-89.
Fukaya, T., Iwakawa, H.O., and Tomari, Y. (2014). MicroRNAs block assembly of eIF4F translation initiation complex in Drosophila. Mol Cell *56*, 67-78.
Galban, S., Kuwano, Y., Pullmann, R., Jr., Martindale, J.L., Kim, H.H., Lal, A., Abdelmohsen, K., Yang, X., Dang, Y., Liu, J.O*., et al.* (2008). RNA-binding proteins HuR and PTB promote the translation of hypoxia-inducible factor 1alpha. Mol Cell Biol *28*, 93-107.
Gebauer, F., Grskovic, M., and Hentze, M.W. (2003). Drosophila sex-lethal inhibits the stable association of the 40S ribosomal subunit with msl-2 mRNA. Mol Cell *11*, 1397-1404.
Gebauer, F., Preiss, T., and Hentze, M.W. (2012). From cis-regulatory elements to complex RNPs and back. Cold Spring Harb Perspect Biol *4*, a012245.

Gherzi, R., Lee, K.Y., Briata, P., Wegmuller, D., Moroni, C., Karin, M., and Chen, C.Y. (2004). A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. Mol Cell *14*, 571-583.
Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods *6*, 343-345.
Gray, N.K., and Hentze, M.W. (1994). Iron regulatory protein prevents binding of the 43S translation pre-initiation complex to ferritin and eALAS mRNAs. EMBO J *13*, 3882-3891.
Grosshans, H., Johnson, T., Reinert, K.L., Gerstein, M., and Slack, F.J. (2005). The temporal patterning microRNA let-7 regulates several transcription factors at the larval to adult transition in C. elegans. Dev Cell *8*, 321-330.
Grskovic, M., Hentze, M.W., and Gebauer, F. (2003). A co-repressor assembly nucleated by Sex-lethal in the 3'UTR mediates translational control of Drosophila msl-2 mRNA. EMBO J *22*, 5571-5581.
Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature *466*, 835-840.
Hafner, M., Max, K.E., Bandaru, P., Morozov, P., Gerstberger, S., Brown, M., Molina, H., and Tuschl, T. (2013). Identification of mRNAs bound and regulated by human LIN28 proteins and molecular requirements for RNA recognition. RNA *19*, 613-626.
Harris, D.T., and Horvitz, H.R. (2011). MAB-10/NAB acts with LIN-29/EGR to regulate terminal differentiation and the transition from larva to adult in C. elegans. Development *138*, 4051-4062.
Hennig, J., Militti, C., Popowicz, G.M., Wang, I., Sonntag, M., Geerlof, A., Gabel, F., Gebauer, F., and Sattler, M. (2014). Structural basis for the assembly of the Sxl-Unr translation regulatory complex. Nature *515*, 287-290.
Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol Cell *32*, 276-284.
Hinman, M.N., and Lou, H. (2008). Diverse molecular functions of Hu proteins. Cell Mol Life Sci *65*, 3168-3181.
Hinnebusch, A.G. (2005). Translational regulation of GCN4 and the general amino acid control of yeast. Annu Rev Microbiol *59*, 407-450.
Hunter, S.E., Finnegan, E.F., Zisoulis, D.G., Lovci, M.T., Melnik-Martinez, K.V., Yeo, G.W., and Pasquinelli, A.E. (2013). Functional genomic analysis of the let-7 regulatory network in Caenorhabditis elegans. PLoS Genet *9*, e1003353.
Hussey, G.S., Chaudhury, A., Dawson, A.E., Lindner, D.J., Knudsen, C.R., Wilce, M.C., Merrick, W.C., and Howe, P.H. (2011). Identification of an mRNP complex regulating tumorigenesis at the translational elongation step. Mol Cell *41*, 419-431.
Ikeda, K., and Inoue, S. (2012). TRIM proteins as RING finger E3 ubiquitin ligases. Adv Exp Med Biol *770*, 27-37.
Inoue, T., Wang, M., Ririe, T.O., Fernandes, J.S., and Sternberg, P.W. (2005). Transcriptional network underlying Caenorhabditis elegans vulval development. Proc Natl Acad Sci U S A *102*, 4972-4977.
Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol *11*, 113-127.
Jänicke, A., Vancuylenberg, J., Boag, P.R., Traven, A., and Beilharz, T.H. (2012). ePAT: a simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. RNA *18*, 1289-1295.
Jin, J., Jing, W., Lei, X.X., Feng, C., Peng, S., Boris-Lawrie, K., and Huang, Y. (2011). Evidence that Lin28 stimulates translation by recruiting RNA helicase A to polysomes. Nucleic Acids Res *39*, 3724-3734.
Johnstone, O., and Lasko, P. (2001). Translational regulation and RNA localization in Drosophila oocytes and embryos. Annu Rev Genet *35*, 365-406.
Jonas, S., and Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. Nat Rev Genet *16*, 421-433.
Jones, K.A. (2007). Transcription strategies in terminally differentiated cells: shaken to the core. Genes Dev *21*, 2113-2117.
Kapasi, P., Chaudhuri, S., Vyas, K., Baus, D., Komar, A.A., Fox, P.L., Merrick, W.C., and Mazumder, B. (2007). L13a blocks 48S assembly: role of a general initiation factor in mRNA-specific translational control. Mol Cell *25*, 113-126.
Katic, I., and Grosshans, H. (2013). Targeted heritable mutation and gene conversion by Cas9-CRISPR in Caenorhabditis elegans. Genetics *195*, 1173-1176.
Katic, I., Xu, L., and Ciosk, R. (2015). CRISPR/Cas9 Genome Editing in Caenorhabditis elegans: Evaluation of Templates for Homology-Mediated Repair and Knock-Ins by Homology-Independent DNA Repair. G3 (Bethesda) *5*, 1649-1656.
Kedde, M., Strasser, M.J., Boldajipour, B., Oude Vrielink, J.A., Slanchev, K., le Sage, C., Nagel, R., Voorhoeve, P.M., van Duijse, J., Orom, U.A*., et al.* (2007). RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. Cell *131*, 1273-1286.
Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J.A., Elkon, R., and Agami, R. (2010). A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. Nat Cell Biol *12*, 1014-1020.
Kelley, R.L., Wang, J., Bell, L., and Kuroda, M.I. (1997). Sex lethal controls dosage compensation in Drosophila by a non-splicing mechanism. Nature *387*, 195-199.
Kimble, J., and Hirsh, D. (1979). The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans. Dev Biol *70*, 396-417.
Koh, K., and Rothman, J.H. (2001). ELT-5 and ELT-6 are required continuously to regulate epidermal seam cell differentiation and cell fusion in C. elegans. Development *128*, 2867-2880.
Kopp, A. (2012). Dmrt genes in the development and evolution of sexual dimorphism. Trends Genet *28*, 175-184.

Kuersten, S., and Goodwin, E.B. (2003). The power of the 3' UTR: translational control and development. Nat Rev Genet *4*, 626-637.
Kühn, L.C. (2015). Iron regulatory proteins and their role in controlling iron metabolism. Metallomics *7*, 232-243.
Kullmann, M., Gopfert, U., Siewe, B., and Hengst, L. (2002). ELAV/Hu proteins inhibit p27 translation via an IRES element in the p27 5'UTR. Genes Dev *16*, 3087-3099.
Kwon, S.C., Yi, H., Eichelbaum, K., Fohr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., and Kim, V.N. (2013). The RNA-binding protein repertoire of embryonic stem cells. Nat Struct Mol Biol *20*, 1122-1130.
Laslo, P., Spooner, C.J., Warmflash, A., Lancki, D.W., Lee, H.J., Sciammas, R., Gantner, B.N., Dinner, A.R., and Singh, H. (2006). Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. Cell *126*, 755-766.
Le, N., Nagarajan, R., Wang, J.Y., Svaren, J., LaPash, C., Araki, T., Schmidt, R.E., and Milbrandt, J. (2005). Nab proteins are essential for peripheral nervous system myelination. Nat Neurosci *8*, 932-940.
Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.
Leeb, M., Dietmann, S., Paramor, M., Niwa, H., and Smith, A. (2014). Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. Cell Stem Cell *14*, 385-393.
Lehrbach, N.J., Armisen, J., Lightfoot, H.L., Murfitt, K.J., Bugaut, A., Balasubramanian, S., and Miska, E.A. (2009). LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. Nat Struct Mol Biol *16*, 1016-1020.
Li, N., Zhong, X., Lin, X., Guo, J., Zou, L., Tanyi, J.L., Shao, Z., Liang, S., Wang, L.P., Hwang, W.T.*, et al.* (2012). Lin-28 homologue A (LIN28A) promotes cell cycle progression via regulation of cyclin-dependent kinase 2 (CDK2), cyclin D1 (CCND1), and cell division cycle 25 homolog A (CDC25A) expression in cancer. J Biol Chem *287*, 17386-17397.
Lin, S.Y., Johnson, S.M., Abraham, M., Vella, M.C., Pasquinelli, A., Gamberi, C., Gottlieb, E., and Slack, F.J. (2003). The C elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. Dev Cell *4*, 639-650.
Loedige, I., Gaidatzis, D., Sack, R., Meister, G., and Filipowicz, W. (2013). The mammalian TRIM-NHL protein TRIM71/LIN-41 is a repressor of mRNA function. Nucleic Acids Res *41*, 518-532.
Loedige, I., Jakob, L., Treiber, T., Ray, D., Stotz, M., Treiber, N., Hennig, J., Cook, K.B., Morris, Q., Hughes, T.R.*, et al.* (2015). The Crystal Structure of the NHL Domain in Complex with RNA Reveals the Molecular Basis of Drosophila Brain-Tumor-Mediated Gene Regulation. Cell Rep *13*, 1206-1220.
Loedige, I., Stotz, M., Qamar, S., Kramer, K., Hennig, J., Schubert, T., Loffler, P., Langst, G., Merkl, R., Urlaub, H.*, et al.* (2014). The NHL domain of BRAT is an RNA-binding domain that directly contacts the hunchback mRNA for regulation. Genes Dev *28*, 749-764.
Loer, B., Bauer, R., Bornheim, R., Grell, J., Kremmer, E., Kolanus, W., and Hoch, M. (2008). The NHL-domain protein Wech is crucial for the integrin-cytoskeleton link. Nat Cell Biol *10*, 422-428.
Lorenz, R., Bernhart, S.H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol Biol *6*, 26.
Mason, D.A., Rabinowitz, J.S., and Portman, D.S. (2008). dmd-3, a doublesex-related gene regulated by tra-1, governs sex-specific morphogenesis in C. elegans. Development *135*, 2373-2382.
Mathys, H., Basquin, J., Ozgur, S., Czarnocki-Cieciura, M., Bonneau, F., Aartse, A., Dziembowski, A., Nowotny, M., Conti, E., and Filipowicz, W. (2014). Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression. Mol Cell *54*, 751-765.
McHugh, C.A., Chen, C.K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A.*, et al.* (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature *521*, 232-236.
Medenbach, J., Seiler, M., and Hentze, M.W. (2011). Translational control via protein-regulated upstream open reading frames. Cell *145*, 902-913.
Meijer, H.A., Kong, Y.W., Lu, W.T., Wilczynska, A., Spriggs, R.V., Robinson, S.W., Godfrey, J.D., Willis, A.E., and Bushell, M. (2013). Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation. Science *340*, 82-85.
Meneely, P.M., McGovern, O.L., Heinis, F.I., and Yanowitz, J.L. (2012). Crossover distribution and frequency are regulated by him-5 in Caenorhabditis elegans. Genetics *190*, 1251-1266.
Merritt, C., Rasoloson, D., Ko, D., and Seydoux, G. (2008). 3' UTRs are the primary regulators of gene expression in the C. elegans germline. Curr Biol *18*, 1476-1482.
Miller, M.A., and Olivas, W.M. (2011). Roles of Puf proteins in mRNA degradation and translation. Wiley Interdiscip Rev RNA *2*, 471-492.
Min, I.M., Pietramaggiori, G., Kim, F.S., Passegue, E., Stevenson, K.E., and Wagers, A.J. (2008). The transcription factor EGR1 controls both the proliferation and localization of hematopoietic stem cells. Cell Stem Cell *2*, 380-391.
Mitchell, S.F., Jain, S., She, M., and Parker, R. (2013). Global analysis of yeast mRNPs. Nat Struct Mol Biol *20*, 127-133.
Mitschka, S., Ulas, T., Goller, T., Schneider, K., Egert, A., Mertens, J., Brustle, O., Schorle, H., Beyer, M., Klee, K.*, et al.* (2015). Co-existence of intact stemness and priming of neural differentiation programs in mES cells lacking Trim71. Sci Rep *5*, 11126.
Moss, E.G. (2007). Heterochronic genes and the nature of developmental time. Curr Biol *17*, R425-434.

Moss, E.G., Lee, R.C., and Ambros, V. (1997). The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. Cell *88*, 637-646.
Nakamura, A., Sato, K., and Hanyu-Nakamura, K. (2004). Drosophila cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA translation in oogenesis. Dev Cell *6*, 69-78.
Newman, A.P., Inoue, T., Wang, M., and Sternberg, P.W. (2000). The Caenorhabditis elegans heterochronic gene lin-29 coordinates the vulval-uterine-epidermal connections. Curr Biol *10*, 1479-1488.
Newman, M.A., Thomson, J.M., and Hammond, S.M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA *14*, 1539-1549.
Nguyen, H.Q., Hoffman-Liebermann, B., and Liebermann, D.A. (1993). The zinc finger transcription factor Egr-1 is essential for and restricts differentiation along the macrophage lineage. Cell *72*, 197-209.
Ostareck, D.H., Ostareck-Lederer, A., Wilm, M., Thiele, B.J., Mann, M., and Hentze, M.W. (1997). mRNA silencing in erythroid differentiation: hnRNP K and hnRNP E1 regulate 15-lipoxygenase translation from the 3' end. Cell *89*, 597-606.
Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degnan, B., Muller, P.*, et al.* (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature *408*, 86-89.
Peng, S., Chen, L.L., Lei, X.X., Yang, L., Lin, H., Carmichael, G.G., and Huang, Y. (2011). Genome-wide studies reveal that Lin28 enhances the translation of genes important for growth and survival of human embryonic stem cells. Stem Cells *29*, 496-504.
Perez-Ortin, J.E., Alepuz, P., Chavez, S., and Choder, M. (2013). Eukaryotic mRNA decay: methodologies, pathways, and links to other stages of gene expression. J Mol Biol *425*, 3750-3775.
Pichon, X., Wilson, L.A., Stoneley, M., Bastide, A., King, H.A., Somers, J., and Willis, A.E. (2012). RNA binding protein/RNA element interactions and the control of translation. Curr Protein Pept Sci *13*, 294-304.
Polesskaya, A., Cuvellier, S., Naguibneva, I., Duquet, A., Moss, E.G., and Harel-Bellan, A. (2007). Lin-28 binds IGF-2 mRNA and participates in skeletal myogenesis by increasing translation efficiency. Genes Dev *21*, 1125-1138.
Ray, D., Kazan, H., Chan, E.T., Pena Castillo, L., Chaudhry, S., Talukder, S., Blencowe, B.J., Morris, Q., and Hughes, T.R. (2009). Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. Nat Biotechnol *27*, 667-670.
Reddy, K.B. (2015). MicroRNA (miRNA) in cancer. Cancer Cell Int *15*, 38.
Rehfeld, F., Rohde, A.M., Nguyen, D.T., and Wulczyn, F.G. (2015). Lin28 and let-7: ancient milestones on the road from pluripotency to neurogenesis. Cell Tissue Res *359*, 145-160.
Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature *403*, 901-906.
Rougvie, A.E., and Ambros, V. (1995). The heterochronic gene lin-29 encodes a zinc finger protein that controls a terminal differentiation event in Caenorhabditis elegans. Development *121*, 2491-2500.
Rougvie, A.E., and Moss, E.G. (2013). Developmental transitions in C. elegans larval stages. Curr Top Dev Biol *105*, 153-180.
Rüegger, S., Miki, T.S., Hess, D., and Grosshans, H. (2015). The ribonucleotidyl transferase USIP-1 acts with SART3 to promote U6 snRNA recycling. Nucleic Acids Res *43*, 3344-3357.
Rybak, A., Fuchs, H., Hadian, K., Smirnova, L., Wulczyn, E.A., Michel, G., Nitsch, R., Krappmann, D., and Wulczyn, F.G. (2009). The let-7 target gene mouse lin-41 is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. Nat Cell Biol *11*, 1411-1420.
Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E.E., Nitsch, R., and Wulczyn, F.G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nat Cell Biol *10*, 987-993.
Sanduja, S., Blanco, F.F., and Dixon, D.A. (2011). The roles of TTP and BRF proteins in regulated mRNA decay. Wiley Interdiscip Rev RNA *2*, 42-57.
Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B.*, et al.* (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676-682.
Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337-342.
Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2013). Corrigendum: Global quantification of mammalian gene expression control. Nature *495*, 126-127.
Schweingruber, C., Rufener, S.C., Zund, D., Yamashita, A., and Muhlemann, O. (2013). Nonsense-mediated mRNA decay - mechanisms of substrate mRNA recognition and degradation in mammalian cells. Biochim Biophys Acta *1829*, 612-623.
Sevetson, B.R., Svaren, J., and Milbrandt, J. (2000). A novel activation function for NAB proteins in EGR-dependent transcription of the luteinizing hormone beta gene. J Biol Chem *275*, 9749-9757.
Shenoy, A., and Blelloch, R.H. (2014). Regulation of microRNA function in somatic stem cell proliferation and differentiation. Nat Rev Mol Cell Biol *15*, 565-576.
Shyh-Chang, N., and Daley, G.Q. (2013). Lin28: primal regulator of growth and metabolism in stem cells. Cell Stem Cell *12*, 395-406.
Siwaszek, A., Ukleja, M., and Dziembowski, A. (2014). Proteins involved in the degradation of cytoplasmic mRNA in the major eukaryotic model systems. RNA Biol *11*, 1122-1136.

Slack, F.J., Basson, M., Liu, Z., Ambros, V., Horvitz, H.R., and Ruvkun, G. (2000). The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. Mol Cell *5*, 659-669.
Spike, C.A., Coetzee, D., Eichten, C., Wang, X., Hansen, D., and 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*, 1535-1558.
Stebbins-Boaz, B., Cao, Q., de Moor, C.H., Mendez, R., and Richter, J.D. (1999). Maskin is a CPEB-associated factor that transiently interacts with eIF-4E. Mol Cell *4*, 1017-1027.
Stiernagle, T. (2006). Maintenance of C. elegans. WormBook, ed. The C. elegans Research Community, doi/10.1895/wormbook.1.101.1.
Stripecke, R., and Hentze, M.W. (1992). Bacteriophage and spliceosomal proteins function as position-dependent cis/trans repressors of mRNA translation in vitro. Nucleic Acids Res *20*, 5555-5564.
Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., and Bartel, D.P. (2014). Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature *508*, 66-71.
Sulston, J.E., and Horvitz, H.R. (1977). Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. Dev Biol *56*, 110-156.
Svaren, J., Sevetson, B.R., Golda, T., Stanton, J.J., Swirnoff, A.H., and Milbrandt, J. (1998). Novel mutants of NAB corepressors enhance activation by Egr transactivators. EMBO J *17*, 6010-6019.
Szostak, E., and Gebauer, F. (2013). Translational control by 3'-UTR-binding proteins. Brief Funct Genomics *12*, 58-65.
Tan, F.E., and Elowitz, M.B. (2014). Brf1 posttranscriptionally regulates pluripotency and differentiation responses downstream of Erk MAP kinase. Proc Natl Acad Sci U S A *111*, E1740-1748.
Tocchini, C., Keusch, J.J., Miller, S.B., Finger, S., Gut, H., Stadler, M.B., and Ciosk, R. (2014). The TRIM-NHL protein LIN-41 controls the onset of developmental plasticity in Caenorhabditis elegans. PLoS Genet *10*, e1004533.
Topilko, P., Schneider-Maunoury, S., Levi, G., Baron-Van Evercooren, A., Chennoufi, A.B., Seitanidou, T., Babinet, C., and Charnay, P. (1994). Krox-20 controls myelination in the peripheral nervous system. Nature *371*, 796-799.
Vadla, B., Kemper, K., Alaimo, J., Heine, C., and Moss, E.G. (2012). lin-28 controls the succession of cell fate choices via two distinct activities. PLoS Genet *8*, e1002588.
Van Wynsberghe, P.M., Kai, Z.S., Massirer, K.B., Burton, V.H., Yeo, G.W., and Pasquinelli, A.E. (2011). LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in Caenorhabditis elegans. Nat Struct Mol Biol *18*, 302-308.
Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science *320*, 97-100.
Volz, K. (2008). The functional duality of iron regulatory protein 1. Curr Opin Struct Biol *18*, 106-111.
White, E.J., Brewer, G., and Wilson, G.M. (2013). Post-transcriptional control of gene expression by AUF1: mechanisms, physiological targets, and regulation. Biochim Biophys Acta *1829*, 680-688.
Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862.
Wilbert, M.L., Huelga, S.C., Kapeli, K., Stark, T.J., Liang, T.Y., Chen, S.X., Yan, B.Y., Nathanson, J.L., Hutt, K.R., Lovci, M.T*., et al.* (2012). LIN28 binds messenger RNAs at GGAGA motifs and regulates splicing factor abundance. Mol Cell *48*, 195-206.
Wilkie, G.S., Dickson, K.S., and Gray, N.K. (2003). Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. Trends Biochem Sci *28*, 182-188.
Worringer, K.A., Rand, T.A., Hayashi, Y., Sami, S., Takahashi, K., Tanabe, K., Narita, M., Srivastava, D., and Yamanaka, S. (2014). The let-7/LIN-41 pathway regulates reprogramming to human induced pluripotent stem cells by controlling expression of prodifferentiation genes. Cell Stem Cell *14*, 40-52.
Wright, J.E., Gaidatzis, D., Senften, M., Farley, B.M., Westhof, E., Ryder, S.P., and Ciosk, R. (2011). A quantitative RNA code for mRNA target selection by the germline fate determinant GLD-1. EMBO J *30*, 533-545.
Wu, X., and Brewer, G. (2012). The regulation of mRNA stability in mammalian cells: 2.0. Gene *500*, 10-21.
Xu, B., Zhang, K., and Huang, Y. (2009). Lin28 modulates cell growth and associates with a subset of cell cycle regulator mRNAs in mouse embryonic stem cells. RNA *15*, 357-361.
Ye, J., and Blelloch, R. (2014). Regulation of pluripotency by RNA binding proteins. Cell Stem Cell *15*, 271-280.
Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R*., et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-1920.
Zhang, D.L., Hughes, R.M., Ollivierre-Wilson, H., Ghosh, M.C., and Rouault, T.A. (2009). A ferroportin transcript that lacks an iron-responsive element enables duodenal and erythroid precursor cells to evade translational repression. Cell Metab *9*, 461-473.
Zou, Y., Chiu, H., Zinovyeva, A., Ambros, V., Chuang, C.F., and Chang, C. (2013). Developmental decline in neuronal regeneration by the progressive change of two intrinsic timers. Science *340*, 372-376.

7 Acknowledgements

I would like to thank my mentor Helge for all his support during the last five years, for his very good ideas and input, his excellent guidance, his enthusiasm for science, for giving me the freedom to develop and realize my own ideas, and most of all for always being positive and motivating.

I am very thankful to Monika, who introduced me to the lab and into the world of *C. elegans*, for always being there for the whole lab when problems occurred and for supporting me in my project.

I would like to thank Iskra and Lene for their immense support, work-related and personal, and for sharing the saddest moments during my time at the FMI with me.

I am very grateful to Gert-Jan for sharing my first PhD project with me, including the endless days (and sometimes nights) at the microscope. Although none of these experiments made it into this thesis, this was a fun project and collaboration.

I would like to thank Lan for all her help with cloning and injections.

I am very thankful to Mihaela Zavolan, Oliver Mühlemann and Marc Bühler, for having joined my thesis committee and for their support, comments and discussions during the meetings.

I would like to thank Lene, Andi, Dimos, Pooja, Hrishi, Rafal, Jerry and Christoph for all the fruitful collaborations, Jun and Chiara for helpful comments on this thesis, Manuel for the routine Tuesday and Thursday tennis sessions, the great facility members at the FMI for all their support, the crew of the media kitchen for being incredibly efficient and reliable and all the current and past members of the Großhans lab for the awesome work atmosphere, for all their help and for all the fun times outside the lab.

And last, but certainly not least, I would like to thank my girlfriend Valérie for all her incredible support, her interest in and her understanding for my peculiar profession; my mother, my sister and my grandparents, who have all supported me in every step of my career and studies; my flat mate Manuel for sharing the home office to write our PhD theses and my other flat mates for taking care of me during the intense months when writing this thesis.