RNA-binding proteins

Min'Ho Lee
Washington University School of Medicine in St. Louis

Tim Schedl
Washington University School of Medicine in St. Louis

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

Recommended Citation
https://digitalcommons.wustl.edu/open_access_pubs/3482

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
RNA-binding proteins*

Min-Ho Lee§†, Tim Schedl§, Department of Genetics, Washington University School of Medicine, St. Louis, MO 63110 USA

Table of Contents

1. Overview ............................................................................................................................... 2
2. RBPs function during germline and early embryo development ........................................................ 5
3. RBPs function in somatic development .................................................................................... 6
4. RNA targets of RBPs ................................................................................................................ 6
5. RNA binding specificity of RBPs ................................................................................................ 7
6. Closing remarks .................................................................................................................... 8
7. Acknowledgements .................................................................................................................. 8
8. References .............................................................................................................................. 8

Abstract

The *C. elegans* genome encodes many RNA-binding proteins (RBPs) with diverse functions in development, indicative of extensive layers of post-transcriptional control of RNA metabolism. A number of *C. elegans* RBPs have been identified by forward or reverse genetics. They tend to display tissue-specific mutant phenotypes, which underscore their functional importance. In addition, several RBPs that bind regulatory sequences in the 3‘untranslated regions of mRNAs have been identified molecularly. Most *C. elegans* RBPs are conserved throughout evolution, suggesting that their study in *C. elegans* may uncover new conserved biological functions. In this review, we primarily discuss RBPs that are associated with well-characterized mutant phenotypes in the germ line, the early embryo, or in somatic tissues. We also discuss the identification of RNA targets of RBPs, which is an important first step to understand how an RBP controls *C. elegans* development. It is likely that most RBPs regulate multiple RNA targets. Once multiple RNA targets are identified, specific features that distinguish target from non-target RNAs and the type(s) of RNA metabolism that each RBP controls can be determined. Furthermore, one can determine whether the RBP regulates all targets by the same mechanism or different targets by distinct mechanisms. Such studies will provide insights into how RBPs exert coordinate control of their RNA targets, thereby affecting development in a concerted fashion.
1. Overview

RNA-binding proteins (RBPs) play key roles in post-transcriptional control of RNAs, which, along with transcriptional regulation, is a major way to regulate patterns of gene expression during development. Post-transcriptional control can occur at many different steps in RNA metabolism, including splicing, polyadenylation, mRNA stability, mRNA localization and translation (Curtis et al., 1995; Wickens et al., 2000; de Moor and Richter, 2001; Johnstone and Lasko, 2001). In the *C. elegans* genome, approximately 500 genes are annotated to encode RBPs (Wormbase), as they have one or more known RNA binding domains such as the RNA Recognition Motif (RRM, also known as RBD or RNP domain), K Homology (KH) domain, Zinc finger (mainly C-x8-C-x5-C-x3-H type), RGG box, DEAD/DEAH box, Pumilio/FBF (PUF) domain, double-stranded RNA binding domain (DS-RBD), Piwi/Argonaute/Zwille (PAZ) domain, Sm domain, etc. Many RBPs have one or more copies of the same RNA binding domain while others have two or more distinct domains. Several RNA binding domains are suggestive for the molecular function of the RBP; DEAD/DEAH box for RNA helicase activity, PAZ domain for short single-stranded RNA binding in RNAi or microRNAs (miRNA) processes, and Sm domain for snoRNA binding in splicing and possibly in tRNA processing. However, other domains only predict RNA binding and do not specifically indicate in which aspect of RNA metabolism they may participate. The known function of RBP homologues in other species can provide insights into RBP function in *C. elegans*. In addition, functional studies of the *C. elegans* RBPs may reveal unexpected roles for the conserved RBPs. For example, mammalian Y14 is a component of the exon-junction complex that mediates nonsense-mediated mRNA decay (NMD; Kataoka et al., 2000; Fribourg et al., 2003; Singh and Lykke-Andersen, 2003). The *C. elegans* homologue of Y14, RPN-4, does not appear to mediate NMD, even though it is still preferentially associated with spliced mRNAs. Interestingly, *C. elegans* Y14, RPN-4, controls germline sex, suggesting that RPN-4 likely has other functions (Kawano et al., 2004).

In *C. elegans*, many RBPs have been identified genetically, which have tissue-specific mutant phenotypes caused by mutations in RBP genes. In addition, several RBPs that bind regulatory sequences in the 3′untranslated regions (3′UTRs) of mRNAs have been identified by molecular methods (Zhang et al., 1997; Jan et al., 1999; Marin and Evans, 2003; Mootz et al., 2004). Even though the mechanisms by which RBPs influence protein expression patterns in their respective tissues are still poorly understood, the association of many RBPs with mutant phenotypes underscores their importance in *C. elegans* development (Table 1). In this review, we will primarily limit our discussion to RBPs that are associated with well-characterized mutant phenotypes.

Many RBPs have been identified as essential factors during germline and early embryo development and RBPs with essential functions in the development of somatic tissues, including neuron, muscle, hypodermis, and excretory cells, as well as in the timing of development have been also identified (see Table 1). These findings indicate that RBPs and post-transcriptional control are employed in most aspects of development. The recent emergence of RBPs implicated in RNAi and miRNA processes, such as DCR-1, ALG-1/-2, PPW-1, RDE-1, and RDE-4, further emphasizing the importance of RBPs and the complexity of post-transcriptional control in *C. elegans* development (Ambros, 2003; Bartel, 2004).

Even though many RBPs have been identified as having critical roles during the development of various tissues in *C. elegans*, it is largely unclear how such RBPs control development, primarily due to the difficulty in identifying their RNA targets. Most RBPs likely have multiple RNA targets. Therefore, their mutant phenotypes may result from the mis-regulation of many RNA targets, which makes it difficult to identify individual RNA targets using classical genetic approaches. Nevertheless, it is essential to identify a majority of the RNA targets of each RBP in order to fully understand the function of the RBP. Furthermore, the identification of RBPs that bind to known regulatory sequences, such as those defined by mutations, is crucial to understanding the mechanism of post-transcriptional control mediated by such sequences.

---
Table 1. RNA-binding proteins associated with characterized mutant phenotypes in *C. elegans*.

<table>
<thead>
<tr>
<th>RBP gene</th>
<th>RNA binding domain</th>
<th>Developmental function</th>
<th>Post-transcriptional function</th>
<th>RNA targets†</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RBP's function during germline and early embryo development</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>gld-1</em></td>
<td>Maxi-KH</td>
<td>Female meiotic prophase progression, germline sex determination, entry of meiotic development, germ cell apoptosis</td>
<td>Translational repressor</td>
<td>Multiple</td>
<td>Francis et al., 1995a; Francis et al., 1995b; Jones and Schedl, 1995; Jan et al., 1999; Clifford et al., 2000; Lee and Schedl, 2001; 2004; Xu et al., 2001; Marin and Evans, 2003; Mootz et al., 2004</td>
</tr>
<tr>
<td><em>gld-3</em></td>
<td>KH-like</td>
<td>Gemline sex determination, germline survival, embryo development</td>
<td>Translational activator?</td>
<td>Unknown</td>
<td>Eckmann et al., 2002; Wang et al., 2002</td>
</tr>
<tr>
<td><em>fbf-1/-2</em></td>
<td>PUF</td>
<td>Gemline sex determination, stem cell maintenance</td>
<td>Translational repressor</td>
<td><em>fmr-3, gld-1, gld-3s, fbf-1, fbf-2</em></td>
<td>Zhang et al., 1997; Crittenden et al., 2002; Eckmann et al., 2004; Lamont et al., 2004</td>
</tr>
<tr>
<td><em>puf-3 to puf-10</em></td>
<td>PUF</td>
<td>Meiotic division of primary spermatocytes (<em>puf-8</em>)</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Lee and Schedl, 2001; Wickens et al., 2001; Subramaniam and Seydoux, 2003</td>
</tr>
<tr>
<td><em>nos-1/-2</em></td>
<td>Nanos RNA binding</td>
<td>Germline survival (with <em>nos-3</em>?)</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Kraemer et al., 1999; Subramaniam and Seydoux, 1999</td>
</tr>
<tr>
<td><em>nos-3</em></td>
<td>Nanos RNA binding</td>
<td>Gemline sex determination, entry of meiotic development</td>
<td>Translational activator?</td>
<td><em>gld-1</em>?</td>
<td>Kraemer et al., 1999; Hansen et al., 2004</td>
</tr>
<tr>
<td><em>daz-1</em></td>
<td>RRM</td>
<td>Female meiotic prophase progression</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Karashima et al., 2000</td>
</tr>
<tr>
<td><em>oma-1/-2</em> (moe-1/-2)</td>
<td>Zn-finger (C-x8-C-x5-C-x3-H type)</td>
<td>Oocyte maturation</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Detwiler et al., 2001; Shimada et al., 2002</td>
</tr>
<tr>
<td><em>pgl-1</em></td>
<td>RGG box</td>
<td>P-Granule maintenance, early germ cell proliferation/differentiation</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Kawasaki et al., 1998</td>
</tr>
<tr>
<td><em>glh-1</em> to <em>glh-4</em></td>
<td>DEAD/DEAH box, Zinc knuckle (CCHC)</td>
<td>Oogenesis and spermatogenesis (<em>glh-1</em> and <em>glh-4</em> redundantly)</td>
<td>RNA helicase</td>
<td>Unknown</td>
<td>Gruidd et al., 1996; Kuznicki et al., 2000</td>
</tr>
<tr>
<td><em>pie-1</em></td>
<td>Zn-finger (C-x8-C-x5-C-x3-H type)</td>
<td>Embryonic germline specification and maintenance</td>
<td>Translational activator?</td>
<td><em>nos-2</em>?</td>
<td>Mello et al., 1992; Mello et al., 1996; Tenenhaus et al., 2001</td>
</tr>
<tr>
<td><em>mex-3</em></td>
<td>KH</td>
<td>Specifying the fate of the germ line</td>
<td>Translational repressor?</td>
<td><em>pal-1</em></td>
<td>Draper et al., 1996; Huang et al., 2002;</td>
</tr>
<tr>
<td>RBP gene</td>
<td>RNA binding domain</td>
<td>Developmental function</td>
<td>Post-transcriptional function</td>
<td>RNA targets</td>
<td>References</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------</td>
<td>---------------------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>-------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td><strong>mex-5/6</strong></td>
<td>Zn-finger</td>
<td>Soma/germline asymmetry in early embryo</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Schubert et al., 2000</td>
</tr>
<tr>
<td><strong>pos-1</strong></td>
<td>Zn-finger</td>
<td>Specifying the fate of the germline blastomere in early embryo</td>
<td>Translational repressor</td>
<td>glp-1</td>
<td>Tabara et al., 1999a; Ogura et al., 2003</td>
</tr>
<tr>
<td><strong>spn-4</strong></td>
<td>RRM</td>
<td>Soma/germline asymmetry in early embryo</td>
<td>Translational activator</td>
<td>glp-1</td>
<td>Gomes et al., 2001; Ogura et al., 2003</td>
</tr>
<tr>
<td><strong>fog-1</strong></td>
<td>RRM, Zn-finger</td>
<td>Gemline sex determination</td>
<td>Cytoplasmic poly(A) element binding?</td>
<td>fog-1?</td>
<td>Luitjens et al., 2000; Jin et al., 2001a; 2001b</td>
</tr>
<tr>
<td><strong>cpb-1</strong></td>
<td>RRM, Zn-finger</td>
<td>Spermatogenesis</td>
<td>Cytoplasmic poly(A) element binding?</td>
<td>Unknown</td>
<td>Luitjens et al., 2000</td>
</tr>
<tr>
<td><strong>mog-1/4/-5</strong></td>
<td>DEAD/DEAH box</td>
<td>Gemline sex determination, embryo development</td>
<td>RNA helicase</td>
<td>Unknown</td>
<td>Puoti and Kimble, 1999; Puoti and Kimble, 2000</td>
</tr>
<tr>
<td><strong>rnp-4</strong></td>
<td>RRM</td>
<td>Gemline sex determination, embryo development</td>
<td>Exon-junction complex?</td>
<td>Unknown</td>
<td>Kawano et al., 2004</td>
</tr>
</tbody>
</table>

**RBP function in RNAi/miRNA processes (germ line and soma)**

<table>
<thead>
<tr>
<th>RBP gene</th>
<th>RNA binding domain</th>
<th>Developmental function</th>
<th>Post-transcriptional function</th>
<th>RNA targets</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>dcr-1</strong></td>
<td>DEAD/DEAH box, DS-RBD, PAZ</td>
<td>Processing of pre-miRNA and long dsRNA, oogenesis</td>
<td>RNase endonuclease</td>
<td>Pre-miRNA, dsRNA</td>
<td>Grishok et al., 2001; Ketting et al., 2001; Knight and Bass, 2001</td>
</tr>
<tr>
<td><strong>alg-1/-2</strong></td>
<td>PAZ</td>
<td>miRNA dependent translational repression and RNAi</td>
<td>RISC complex</td>
<td>Short ssRNA</td>
<td>Grishok et al., 2001</td>
</tr>
<tr>
<td><strong>rde-1</strong></td>
<td>PAZ</td>
<td>miRNA dependent translational repression and RNAi</td>
<td>RISC complex</td>
<td>Short ssRNA</td>
<td>Tabara et al., 1999b; Tabara et al., 2002</td>
</tr>
<tr>
<td><strong>ppw-1</strong></td>
<td>PAZ</td>
<td>Germline RNAi</td>
<td>RISC complex</td>
<td>Short ssRNA</td>
<td>Tijsterman et al., 2002</td>
</tr>
</tbody>
</table>

**RBP function in somatic development**

**RBP function in primary sex determination**

<table>
<thead>
<tr>
<th>RBP gene</th>
<th>RNA binding domain</th>
<th>Developmental function</th>
<th>Post-transcriptional function</th>
<th>RNA targets</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>fox-1</strong></td>
<td>RRM</td>
<td>Somatic sex determination, X-chromosome counting</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Nicoll et al., 1997; Skipper et al., 1999</td>
</tr>
</tbody>
</table>

**RBP function in neuromuscular system**

<table>
<thead>
<tr>
<th>RBP gene</th>
<th>RNA binding domain</th>
<th>Developmental function</th>
<th>Post-transcriptional function</th>
<th>RNA targets</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>unc-75</strong></td>
<td>RRM</td>
<td>Fine tuning of synaptic transmission</td>
<td>Putative splicing factor</td>
<td>Unknown</td>
<td>Loria et al., 2003</td>
</tr>
<tr>
<td><strong>exc-7</strong></td>
<td>RRM</td>
<td>Fine tuning of synaptic transmission</td>
<td>Putative splicing factor</td>
<td>Unknown</td>
<td>Loria et al., 2003</td>
</tr>
<tr>
<td>RBP gene</td>
<td>RNA binding domain</td>
<td>Developmental function</td>
<td>Post-transcriptional function</td>
<td>RNA targets</td>
<td>References</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------</td>
<td>------------------------</td>
<td>-----------------------------</td>
<td>-------------</td>
<td>------------</td>
</tr>
<tr>
<td><em>etr-1</em></td>
<td>RRM</td>
<td>Muscle formation and function</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Milne and Hodgkin, 1999</td>
</tr>
<tr>
<td><em>mec-8</em></td>
<td>RRM</td>
<td>Specific alternative splices of <em>unc-52</em> transcripts in hypodermis</td>
<td>Putative splicing factor</td>
<td><em>unc-52</em>?</td>
<td>Lundquist et al., 1996; Spike et al., 2002</td>
</tr>
<tr>
<td><em>msl-1</em></td>
<td>RRM</td>
<td>Male mating behavior</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Yoda et al., 2000</td>
</tr>
</tbody>
</table>

RBP function in excretory cells

<table>
<thead>
<tr>
<th>RBP gene</th>
<th>RNA binding domain</th>
<th>Developmental function</th>
<th>Post-transcriptional function</th>
<th>RNA targets</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>exc-7</em></td>
<td>RRM</td>
<td>Excretory canals development.</td>
<td>Putative splicing factor</td>
<td><em>sma-1</em>?</td>
<td>Fujita et al., 2003</td>
</tr>
</tbody>
</table>

RBP function as a heterochronic regulator

<table>
<thead>
<tr>
<th>RBP gene</th>
<th>RNA binding domain</th>
<th>Developmental function</th>
<th>Post-transcriptional function</th>
<th>RNA targets</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>lin-28</em></td>
<td>Cold Shock Domain, Zinc knuckle (CCHC)</td>
<td>A regulator of developmental timing in the first larval stage</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Moss et al., 1997; Moss and Tang, 2003</td>
</tr>
</tbody>
</table>

Footnotes

Proteins known to bind RBPs and regulate RNA metabolism such as GLD-2 (Wang et al., 2002; Eckmann et al., 2004) and ATX-2 (Kiehl et al., 2000; Cioks et al., 2004) are not included in this table.

1 Only characterized RNA targets are listed. However, most RBPs likely have additional RNA targets. RNA targets with “?” indicate that the direct relationship between RBPs and RNA targets has not yet been proven.

2 *C. elegans* has ~10 *puf* genes while *Drosophila* and mammals have fewer, suggesting that an expansion in the *puf* family has occurred in *C. elegans*. Each PUF may have unique function while different combinations of PUF proteins have redundant functions, suggesting that each PUF likely has its own RNA targets and some of the targets may be regulated by more than one PUF.

2. RBPs function during germline and early embryo development

Many RBPs have essential functions during late germline and early embryo development (Table 1) because post-transcriptional control of maternal mRNAs is a predominant mechanism for temporal/spatial regulation of gene expression during this period (Wickens et al., 2000; de Moor and Richter, 2001; Kuersten and Goodwin, 2003). In general, RBPs bind regulatory sequences that are usually located in 5'UTR and/or 3'UTR of the mRNAs to exert post-transcriptional regulation. During meiotic prophase progression, chromosomes become condensed at diakinesis, and the genome becomes transcriptionally silent (Gibert et al., 1984; Schisa et al., 2001; Kelly et al., 2002). Therefore, during the process of late oogenesis (and late spermatogenesis), fertilization/meiotic divisions, and early embryogenesis, post-transcriptional control of pre-existing mRNAs, mainly through localization and translational regulation, is the predominant mechanism regulating protein expression. However, post-transcriptional control in the germline also occurs during periods of active transcription (distal mitotic region through the pachytene stage; Crittenden et al., 2002; Marin and Evans, 2003; Hansen et al., 2004; Lee and Schedl, 2001, 2004).

As in other systems, RBP regulatory networks are beginning to be uncovered. For example, FBF-1/-2 are repressors while NOS-3 functions redundantly with a poly-A polymerase, GLD-2, as a putative direct activator of *gld-1* mRNA translation (Crittenden et al., 2002; Hansen et al., 2004). GLD-1 in turn is a putative repressor of *mex-3* mRNA translation and MEX-3 and GLD-1 are spatially non-overlapping repressors of the translation of the *pal-1* mRNA (Mootz et al., 2004). Interestingly, a number of RBPs and some specific mRNAs are localized, at least transiently, to germline specific granules, P granules (Schisa et al., 2001; Barbee et al., 2002 and references therein). However, the function of these associations, particularly in regard to post-transcriptional control, remains to be understood.

Germline sex determination and the proliferation vs. meiotic development decision (stem cell maintenance) rely heavily on post-transcriptional mechanisms in the control of these processes (Table 1). Interestingly, the RBPs, GLD-1, GLD-3, FBF-1/-2, and NOS-3 function in both processes (Francis et al., 1995a; Francis et al., 1995b; Jones...
and Schedl, 1995; Zhang et al., 1997; Kraemer et al., 1999; Crittenden et al., 2002; Eckmann et al., 2002; Hansen et al., 2004). It is unclear at this time whether the observation that these RBPs function in both processes is due to (1) a temporal/spatial regulatory coordination between the sex determination and the proliferation vs. meiosis decisions or (2) the various RBPs having numerous mRNA targets, and the sex determination and the proliferation vs. meiosis phenotypes are just the processes that have been uncovered genetically so far. Other RBPs are known to function in regulating the sex determination decision, such as FOG-1, MOG-1/4/5, and RNP-4 (Ce-Y14) (Puoti et al, 1999, 2000; Jin et al., 2001a; Kawano et al., 2004). It is currently unknown if they also function in the proliferation vs. meiotic development decision.

Several RBPs have been shown to function in meiotic prophase progression, gametogenesis and oocyte maturation such as GLD-1, GLD-3, DAZ-1, OMA-1/-2, PGL-1, and GLH-1/2/3/4. Others (MEX-3, PIE-1, POS-1, GLD-1 and SPN-4) have also been implicated in regulating aspects of early embryo development. These RBPs likely regulate the temporal and spatial pattern of protein expression of subsets of maternal mRNAs, which is crucial in the absence of de novo mRNA synthesis.

3. RBPs function in somatic development

Post-transcriptional control is also important in somatic development as a number of RBPs that have somatic tissue-specific mutant phenotypes have been identified genetically (Table 1). Unlike RBPs that function in germline and early embryo development, several RBPs essential for somatic development are proposed to act as splicing factors, presumably regulating tissue-specific alternative splicing of their mRNA targets. For example, two RRM domain-containing RBPs, MEC-8 and UNC-75, localize to nuclear speckles in the hypodermis and nervous system, respectively (Spike et al., 2002; Loria et al., 2003), consistent with a role in pre-mRNA splicing. Furthermore, MEC-8 regulates the alternative splicing of unc-52 pre-mRNA primarily in the hyperdermis (Spike et al., 2002). Although this regulation is likely direct, it has not yet been proved. In addition, other mec-8 loss-of-function phenotypes are likely independent of unc-52 function, suggesting that MEC-8 regulates additional pre-mRNAs. Another RRM domain-containing RBP, EXC-7, is nuclear localized in embryonic excretory canal cells and throughout the nervous system and is proposed to regulate its mRNA targets through the control of splicing and/or stability (Fujita et al., 2003; Loria et al., 2003).

Somatic RBPs are not limited to regulating RNA processing in the nucleus, as two other somatic RBPs, MSI-1 and LIN-28, are enriched in the cytoplasm (Moss et al., 1997; Yoda et al., 2000). RNAi and miRNA dependent regulation occur during both germline and somatic development and recent discoveries of RBPs functioning in RNAi and miRNA processes, such as DCR-1, ALG-1/2, PPW-1, RDE-1, and RDE-4 (Table 1), has provided a glimpse into understanding this new mode of post-transcriptional control in C. elegans development.

4. RNA targets of RBPs

Several approaches have been taken to identify RNA targets of RBPs or to identify RBPs that bind to known regulatory sequences (Lundquist et al., 1996; Zhang et al., 1997; Jan et al., 1999; Lee and Schedl, 2001; Xu et al., 2001; Fujita et al., 2003; Marin and Evans, 2003; Mootz et al., 2004). In general, genetic analysis indicates that the RBPs must regulate multiple RNA targets because loss of an RBP has more pleiotropic mutant phenotypes than mis-regulations of the known mRNA targets. For example, fbf-1/-2 loss-of-function phenotype suggests that FBF-1/-2 must bind to other mRNAs to maintain germline stem cells in addition to binding fem-3 mRNA to regulate germline sex. Indeed, FBF-1/-2 have been shown to bind and repress the translation of gld-1 mRNA to regulate germline stem cell proliferation (Crittenden et al., 2002). These results support the view that a comprehensive understanding of how such RBPs control development requires the identification of many or all of their mRNA targets. The identification of multiple mRNA targets will allow one to identify sequences and/or structures that distinguish target from non-target RNAs, as well as to determine the type of RNA metabolism the RBP controls.

RNA targets of RBPs have been identified by candidate gene approaches (“educated guess”), as genes that encode putative targets that have similar (or opposite) mutant phenotypes to that of RBP, and/or by studying the expression patterns of the RBPs and the proteins of their putative mRNA targets. For example, 1) Mutations in mec-8 strongly enhance the phenotype of specific mutations in unc-52 and MEC-8 controls alternative splicing of unc-52 transcripts (Lundquist et al., 1996; Spike et al., 2002). 2) exc-7 mutants exhibit synergistic excretory canal defects with mutations in sma-1 and EXC-7 binds to sma-1 mRNA (Fujita et al., 2003). 3) GLP-1 is mis-expressed in pos-1 mutant embryos and POS-1 binds to the glp-1 3’UTR (Ogura et al., 2003). 4) GLD-1 is a cytoplasmic RNA binding protein that represses the translation of several mRNA targets in the distal germline. It was also known that mes-3 and pal-1 mRNAs are translationally repressed in the distal germ line. Therefore it was directly tested and
confirmed that GLD-1 binds and represses the translation of *mes-3* and *pal-1* mRNAs (Xu et al., 2001; Mootz et al., 2004). Because indirect regulation by the RBP also fits the criteria used to predict the candidate mRNA target, these approaches require the experimental test of whether the RBP binds to the candidate mRNA target directly.

Gain-of-function mutations in the *fem-3* and the *tra-2* 3’UTRs (Ahringer and Kimble, 1991; Goodwin et al., 1993), which result in deregulated expression, provided a starting point for molecular identification of RBPs that function in translational repression of these mRNAs. For example, FBF-1/2 were identified by the yeast three-hybrid system using the 3’UTR of *fem-3* mRNA as a bait. FBF-1/2 bind to the 3’UTR of *fem-3* to repress the male sexual fate in the *C. elegans* hermaphrodite germline (Zhang et al., 1997). GLD-1 was found to bind to the *tra-2* 3’UTR, also by using the yeast three-hybrid system (Jan et al., 1999). GLD-1 was also identified as binding to the *glp-1* 3’UTR biochemically (Marin and Evans, 2003).

More direct methods to identify multiple endogenous RNAs associated with RBPs using high-throughput gene array technologies have been described (Tenenbaum et al., 2000, 2002; Brown et al., 2001). They are based on the isolation of endogenous RNA-protein complexes under optimized conditions mostly by immunoprecipitation (IP). The isolated endogenous RNAs are then identified by microarray analysis. This approach has the potential to identify most RNA targets of an RPB without prior knowledge of mutant phenotypes or expression patterns. Mis-regulations of such RNA targets may have small effects or contributions to phenotypes of the RBP. This approach can also uncover mRNA targets from functionally redundant paralagous genes that are co-regulated by an RBP (Lee and Schedl, 2001, 2004).

Sixteen mRNA targets of GLD-1 have been identified by a similar approach (Lee and Schedl, 2001, 2004). Functional GLD-1 was immunoprecipitated from cytosol extracts. The isolated mRNAs were identified after subtractive hybridization/cloning/sequencing. GLD-1 represses the translation of several targets through direct binding (Lee and Schedl, 2001, 2004). Recently, an essentially identical experiment was performed with microarray analysis to detect specifically enriched mRNAs in the GLD-1 IP. This resulted in the identification of significantly more mRNAs (129), which are enriched more than two fold (p < 0.05) in the GLD-1 IP over the control IP (Lee, M.-H., Reinke, V., and Schedl, T., unpubl. data). Essentially, all targets identified previously were identified again. These and other results indicate that most of the 129 mRNA targets are likely to interact with GLD-1 *in vivo* and demonstrate that there is a significant increase in mRNA target detection power with the microarray analysis compared with the subtractive hybridization/cloning/sequencing strategy.

Several efforts are currently undertaken to identify *in vivo* mRNA targets of RBPs, PIE-1 (D’Agostino, I., Reinke, V., and Seydoux, G., pers. comm.), DAZ-1 (Karashima, T., Sugioka, K., Uesugi, K., Kohara, Y., and Yamamoto, M., pers. comm.), and LIN-28 (Moss, E. and Kemper, K., pers. comm.), or to identify *in vivo* mRNA targets of a putative cytoplasmic poly-A polymerase, GLD-2, that interacts with RBPs (Suh, N. and Kimble, J., pers. comm.).

5. RNA binding specificity of RBPs

One of the important questions to understand the function of RBPs in *C. elegans* development is how RBPs distinguish their targets from non-targets *in vivo*. In other words, how RBPs specifically recognize their RNA targets. RBPs may recognize specific sequences, structures, or both, which are present in their RNA targets. Understanding of RNA binding specificity of an RBP can be another way to identify unknown targets that contain similar features, if they have enough information to distinguish targets from non-targets computationally. It will also provide important tools to find the molecular mechanism of the post-transcriptional process that the RBP controls. RNA binding specificity of most RBPs is unknown because their RNA targets are not identified yet. However, RNA binding specificity of GLD-1 and FBF-1/2 in their mRNA targets are beginning to emerge.

GLD-1 has been shown to form a homodimer to bind to two close sites in a single TGE (*tra-2* GLI element) of *tra-2* 3’UTR *in vitro* where one site has a hexanucleotide consensus (Ryder et al., 2004). Among other mRNA targets of GLD-1 identified in the GLD-1 IP, thirteen independent GLD-1 binding regions in seven targets have been found where GLD-1 binds 5’UTR, 3’UTR or in the open reading frame (ORF) depending on the target (Lee and Schedl, 2001, 2004, unpubl. data). Interestingly, many GLD-1 binding regions have the hexanucleotide consensus, suggesting it is likely important for GLD-1 binding. However, several GLD-1 binding regions do not contain the hexanucleotide consensus, suggesting that other distinct features that dictate GLD-1 binding likely exist (Lee, M.-H., and Schedl, T., unpubl. data).
FBF-1/-2 have been shown to bind to sites that carry crucial UGU(G/A) motif in the 3’UTRs of their mRNA targets (Zhang et al., 1997; Crittenden et al., 2002; Eckmann et al., 2004; Lamont et al., 2004). It is interesting to note that FBF-1/2 do not bind to all potential sites that contain UGU(G/A) motif in the 3’UTRs of their mRNA targets (Eckmann et al., 2004; Lamont et al., 2004). This suggests that this motif alone is not sufficient to be recognized by FBF-1/-2 and other distinct features should exist.

6. Closing remarks

The C. elegans genome encodes many RBPs with diverse functions in development, revealing a large layer of post-transcriptional control of RNA metabolism, which had not been previously appreciated in the control of gene expression. Most RBPs with tissue-specific functions are conserved throughout evolution, suggesting that their studies in C. elegans may uncover new conserved biological functions.

In general, RBPs likely regulate multiple RNA targets and the identification of many or all RNA targets of RBPs will be an important first step in a comprehensive understanding of how RBPs control C. elegans development. With multiple RNA targets, specific features that distinguish target from non-target RNAs, as well as the type(s) of RNA metabolism that each RBP controls can be determined. Furthermore, with the identity of multiple RNA targets, one can determine whether the RBP regulates all targets by the same mechanism or different targets by distinct mechanisms. Such studies will provide insights into how RBPs exert coordinate control of their RNA targets, thereby affecting development in a concerted fashion.

7. Acknowledgements

We are grateful to Sudhir Nayak for providing information about genes that are annotated to encode RNA-binding proteins in Wormbase and Dave Hansen and members of the Schedl lab for comments on this review. TS is supported by NIH grant GM63310.

8. References


RNA-binding proteins


