3' UTRs Are the Primary Regulators of Gene Expression in the *C. elegans* Germline

Christopher Merritt, 1 Dominique Rasoloson, 1 Darae Ko, 1 and Geraldine Seydoux 1,*

1Department of Molecular Biology and Genetics
Howard Hughes Medical Institute
Center for Cell Dynamics
Johns Hopkins School of Medicine
725 North Wolfe Street
PCTB 706
Baltimore, Maryland 21205

Summary

How genes are regulated to produce the correct assortment of proteins for every cell type is a fundamental question in biology. For many genes, regulation begins at the DNA level with the use of promoter sequences to control transcription. Regulation can also occur after transcription using sequences in the 3' untranslated region (UTR) of the mRNA to affect mRNA stability and/or translation [1]. The *C. elegans* gonad is an excellent tissue to study gene regulation during development: In the adult, germ cells are arranged in order of differentiation, with undifferentiated progenitors at one end of the gonad, cells in meiotic prophase in the middle, and gametes at the other end [2]. Using a transgenic assay, we have compared the contribution of promoters and 3' UTRs to gene regulation during germline development. We find that for most genes tested, 3' UTRs are sufficient for regulation. With the exception of promoters activated during spermatogenesis, promoters are permissive for expression in all germ cell types (from progenitors to oocytes and sperm). In progenitors, 3' UTRs inhibit the production of meiotic and oocyte proteins by posttranscriptional mechanisms involving PUF- and KH-domain RNA-binding proteins. Our findings indicate that many genes rely primarily on 3' UTRs, not promoters, for regulation during germline development.

Results and Discussion

3' UTRs Are Sufficient to Specify Most Expression Patterns in the Germline

Among genes with known protein distributions in the *C. elegans* germline, in vivo evidence for the involvement of specific cis-regulatory sequences had been obtained for four genes prior to this work. In each case, 3' UTR sequences were found to be necessary ([glp-1, cep-1] or sufficient ([gip-1, pal-1]) for regulation [3–6]. To investigate the extent to which 3' UTRs contribute to gene regulation in the germline, we selected 30 genes representative of every germline expression pattern reported in the literature, including 25 of the 35 genes known to encode proteins expressed in specific germ cell types (Supplemental Experimental Procedures available online). We cloned the 3' UTR of each gene downstream of green fluorescent protein (GFP) fused to Histone H2B and used the *pie-1* promoter to drive all 30 fusions (Figure 1A). The *pie-1* gene encodes a maternal protein expressed specifically in oocytes and embryos, but the *pie-1* promoter permits expression in all germ cell types (reference [7] and see below). Each fusion was randomly integrated in the genome by microparticle bombardment [8]. 30/30 fusions yielded lines that expressed GFP:H2B in the germline, and for 24/30 fusions, the GFP:H2B expression pattern matched that reported for the corresponding endogenous protein (compare nuclear GFP fluorescence to red stippling in Figures 1, 2, and 3A and Figure S1). For example, the *glp-1*, *glld-1*, and *puf-5* 3' UTR fusions (Figure 1C) express GFP:H2B preferentially in the distal, pachytene, and proximal regions, respectively, as has been reported for the GLP-1, GLD-1, and PUF-5 proteins (Table S1). Although all 24 patterns generally matched endogenous patterns, in many cases pattern “boundaries” (areas where GFP:H2B levels rise or decline) appeared expanded compared to those reported in the literature (pattern “expansion” is denoted by hatching in Figure 2). For example, p53 (CEP-1) levels are high in progenitors, low in pachytene, and high in oocytes. The cep-1 3' UTR fusion approximates this pattern, except that GFP:H2B steadily declines through the first half of the pachytene region and reaches minimal levels only in the second half (Figure 1C). Replacement of H2B with the native ORF led to a more accurate pattern for cep-1 and several other fusions (Figure S2), suggesting that the expanded boundaries are a GFP:H2B-specific artifact, perhaps because of slow turnover of the fusion. In total, among the 30 3' UTRs surveyed, we could identify 17 distinct expression patterns (Figure 2), spanning all stages of germ cell development (with one exception, see below). We conclude that 3' UTRs are sufficient to specify a wide range of expression patterns in the germline.

3' UTRs Are Not Sufficient to Specify Sperm-Specific Expression

The only cell-type specificity not reproduced in our survey was sperm-specific expression. Five genes in our survey encode proteins expressed only during spermatogenesis. FOG-1 is expressed in pachytene cells during spermatogenesis and MSP-142, SPE-11, SPE-38, and SPE-41 are expressed in differentiating spermatocytes and mature sperm (Table S1 for references). The *msp-142*, *spe-11*, *spe-38*, and *spe-41* 3' UTR fusions caused GFP:H2B to be expressed in all germ cells, including progenitors, oocytes, and sperm (Figure 2, Figure 3A, and Figure S1). Replacement of H2B with the native ORF (for *msp-142*, *spe-11*, and *spe-38*) did not correct the ubiquitous expression (Figure S2). The *foq-1* 3' UTR fusion showed the expected pachytene enrichment, but remained on in pachytene cells during oogenesis (Figure S1). Downregulation of *foq-1* in oogenic germelines depends on the transcription factor *tra-1* [9], and four consensus *tra-1* binding sites are present in the *foq-1* promoter [10]. We conclude that 3' UTRs are not sufficient to direct sperm-specific expression and that promoters may be required.

Promoters Are Permissive for Expression in All Germ Cells, with the Exception of Promoters Activated during Spermatogenesis

To examine the role of promoters directly, we constructed promoter fusions for 12 genes encoding proteins with restricted
patterns, including *fbf-1, fbf-2, daz-1* (distal germline), *gli-1, lip-1, him-3* (meiotic prophase), *pgl-3, mex-5, pie-1, spn-4* (oocytes), *spe-11*, and *msp-56* (sperm) (Figure 3A and Figure S3). The promoters were fused to GFP:H2B and a 3′ UTR permissive for expression in all germ cells (3′ UTR of *tbb-2*, coding for beta-tubulin) (Figure 1A). Of the 12 promoters tested, only *spe-11* and *msp-56* drove GFP:H2B in the predicted pattern (Figure 3A and Figure S3), confirming that sperm-specific expression depends on promoter sequences (also see [11]). Surprisingly, all other promoter fusions caused GFP:H2B to be expressed in all germ cells. GFP:H2B levels often appeared highest in the pachytene region (most obvious for *spn-4*, Figure 3A). The pachytene region is the most transcriptionally active domain in the adult gonad as determined by rates of tritiated uridine incorporation [12]. Furthermore, functional genomic studies have suggested that genes coding for oocyte proteins, such as *spn-4*, are transcriptionally upregulated during pachytene in adult hermaphrodites [13]. Our results are consistent with those findings but suggest that these genes are also transcribed (at least at basal levels) in other cells, including germline progenitors.

To determine whether promoters coding for oocyte proteins are active in progenitors before the onset of gametogenesis, we examined the expression of five promoter fusions in larval gonads containing only germline progenitors (L2 stage). We examined promoter fusions from three genes encoding oocyte proteins, including *gld-1, lip-1, and mes-5* (Figure 3B). Our results are consistent with those findings but suggest that these genes are also transcribed (at least at basal levels) in other cells, including germline progenitors. (A) Schematics showing the design of 3′ UTR and promoter fusions. Both contain green fluorescent protein (GFP) fused to histone H2B as the protein reporter. H2B associates dynamically with chromatin [30]. Chromatin localization facilitates detection and turnover is fast enough to visualize changes in gene expression even in nondividing cells ([7] and this study). The pie-1 promoter and β-tubulin 3′ UTR (tbb-2) are permissive for expression in all germ cells. (B) Photomicrograph of a single adult gonadal arm (outlined with stippled line) expressing the *mes-2* 3′ UTR fusion (ubiquitous expression). By this stage, hermaphrodites are producing oocytes. Sperm (produced during the L4 larval stage) are stored in the spermatheca. Gonad is divided from distal to proximal into different zones used for systematic scoring of expression patterns as in Figure 2: “A” is the first half of mitotic zone, containing germline stem cells, “B” is the second half of mitotic zone containing cells that have initiated meiotic S phase, “C” and “D” are pachytene (meiotic prophase) regions, “E” is the loop region where cells transition from pachytene to diakinesis and initiate oocyte formation, “F”–“H” is the region of oocyte growth, and “H” denotes last oocyte. (C) Photomicrographs as in (B) showing examples of different 3′ UTR fusions, arranged on the basis of the primary site of protein expression as reported in the literature. For each gene, red stippling outlines the region of the gonad where high levels of endogenous protein have been detected (Table S1, for references). See Figure 3 and Figure S1 for photomicrographs of all other 3′ UTR fusions examined in this study and Figure 2 for a schematic summary. In this and all subsequent figures, vertical lines are used to indicate where two photomicrographs of the same animal have been merged.
proteins (spn-4, mex-5, and pie-1), one gene encoding a sperm protein (spe-11), and one gene encoding a protein enriched in progenitors (fbf-1) (Figure 3B). We detected GFP:H2B expression from the spn-4, mex-5, pie-1, and fbf-1 fusions, but not from the spe-11 fusion. In the case of pie-1, we detected GFP:H2B in gonads containing as few as 8–12 germ cells (but not earlier), indicating that this promoter becomes active after the primordial germ cells begin proliferation in the first larval stage (data not shown). In contrast, expression of the spe-11 fusion was first detected in the L4 stage, in cells at the end of the pachytene region and continuing through mature sperm (Figure 3C). In adult hermaphrodites (which have switched to oogenesis), spe-11:GFP:H2B was no longer detected in pachytene cells but was still present in mature sperm stored in the spermatheca (Figure 3A). In males, which produce sperm continuously, spe-11:GFP:H2B remained expressed in late pachytene cells and later stages of spermatogenesis through adulthood (Figure 3S). In contrast, pie-1:GFP:H2B was detected in all cells in both hermaphrodite and male germ lines (Figure 3S). Therefore, with the exception of sperm-specific genes, the promoters in our survey are not sufficient to specify cell-type specific expression when paired with a permissive 3' UTR such as tbb-2. We conclude that 3' UTRs are the primary source of gene regulation in the germline, with the exception of genes expressed during spermatogenesis. We note, however, that because all promoters were tested in transgenes, we cannot exclude that some promoters do in fact exhibit specificity when in their endogenous chromatin context.

3' UTRs Function Posttranscriptionally and Depend on Two Classes of RNA-Binding Proteins to Inhibit Expression in Progenitors

3' UTR regulatory elements are expected to function after transcription at the mRNA level but in principle could also function

<table>
<thead>
<tr>
<th>Gene</th>
<th>Domain/Homolog</th>
<th>Progenitors</th>
<th>Pachytene</th>
<th>Loop</th>
<th>Oocytes</th>
<th>Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>spn-4</td>
<td>F-box</td>
<td>A</td>
<td>C</td>
<td>D</td>
<td>F</td>
<td>1</td>
</tr>
<tr>
<td>mes-2</td>
<td>Enhancer of Zeste</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>par-5</td>
<td>14-3-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>tbb-2</td>
<td>β-tubulin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>pgl-1</td>
<td>RNA-binding (GG)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>nos-3</td>
<td>RNA-binding (Nanos)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>gsp-1</td>
<td>Notch receptor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>fbf-1</td>
<td>RNA-binding (PUF)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>fbf-2</td>
<td>RNA-binding (PUF)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>pgl-3</td>
<td>RNA-binding (GG)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>gsk-1</td>
<td>RNA-binding (KH)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>hmr-1</td>
<td>Synaptosomal complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>lip-1</td>
<td>MAP kinase phosphatase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>me-2</td>
<td>Yolk (LDL) receptor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>mex-5</td>
<td>CCCH finger</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>puf-5</td>
<td>RNA-binding (PUF)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>psl-1</td>
<td>Transcription factor (Caudal)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>spn-4</td>
<td>RNA-binding (RNP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>pie-1</td>
<td>CCCH finger</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>pos-1</td>
<td>CCCH finger</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>cye-1</td>
<td>Cyclin E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>mes-3</td>
<td>Member of Polycomb complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>ces-1</td>
<td>p53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>mes-3</td>
<td>RNA-binding (KH)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>fog-1</td>
<td>CPEB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>msp-142</td>
<td>Major Sperm Protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>spe-11</td>
<td>Novel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>spe-38</td>
<td>Novel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>spe-41</td>
<td>TRF4 channel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
</tr>
</tbody>
</table>

3' UTRs Function Posttranscriptionally and Depend on Two Classes of RNA-Binding Proteins to Inhibit Expression in Progenitors

3' UTR regulatory elements are expected to function after transcription at the mRNA level but in principle could also function
at the DNA level as distal enhancers to modulate transcription. To distinguish between these possibilities, we tested six 3' UTRs in the context of an "operon" (Figure 4A). C. elegans operons are transcribed into polycistronic primary transcripts, which are processed by transsplicing into monocistronic mRNAs, each with its own 3' UTR and poly(A) tail. We constructed three operon transgenes containing an mCherry fusion linked to a first 3' UTR (gld-1, fbf-2, or him-3) and a GFP fusion linked to a second 3' UTR (cye-1, mex-5, or rme-2). If the 3' UTRs regulate transcription, we would expect the two fusions to be expressed in the same pattern. Instead we found that each reporter behaved independently and was expressed in the pattern expected for its 3' UTR (Figures 4B–4D). We conclude that the 3' UTRs tested function posttranscriptionally to regulate protein expression.

To investigate the mechanisms by which 3' UTRs confer cell-type specificity, we first examined the ability of a "somatic" 3' UTR to support germline expression. The unc-54 gene encodes myosin and is only expressed in muscle. The unc-54 3' UTR is the standard 3' UTR used in transgenes in C. elegans and is permissive for expression in all somatic cells [14]. We found that a 3' UTR fusion containing the pie-1 promoter and the unc-54 3' UTR fusion was expressed in all germ cells (Figure S1). This result suggests that the "default" state in germ cells is on and that 3' UTRs confer specificity primarily by blocking expression in specific cell types (although activating mechanisms may also be needed in the context of germline 3' UTRs [15]).

Several posttranscriptional repressors have been implicated in gene regulation in the germline [16]. The best characterized are two classes of RNA-binding proteins that function in the distal gonad: the PUF domain proteins FBF-1 and FBF-2 and the KH domain proteins MEX-3 and GLD-1. FBF-1 and FBF-2 are nearly identical and are required redundantly for the posttranscriptional repression of several genes in the mitotic zone, including fog-1, gld-1, and lip-1 in our survey.
MEX-3 and GLD-1 are two unrelated KH domain proteins expressed in complementary patterns in the mitotic (MEX-3) and pachytene (GLD-1) regions. MEX-3 and GLD-1 have been implicated in the negative regulation of several mRNAs, including rme-2, pal-1, mex-5, and put-5 in our survey [4, 18, 19]. Consistent with regulation by these proteins, the fog-1, gld-1, lip-1, rme-2, pal-1, mex-5, and put-5 3′ UTR fusions all showed low or no GFP:H2B expression in the distal region of the gonad (Figure 2). To test directly whether these fusions are regulated by FBF-1/2 or MEX-3 and GLD-1 by RNAi. Complete loss of these proteins caused dramatic cell transformations in the adult stage, including loss of all progenitors (fbf-1; fbf-2 double mutant [20]) and transdifferentiation of cells in the pachytene region to somatic fates (gld-1; mex-3 mutant [21]). To minimize secondary effects due to cell fate transformations, we used RNA-mediated interference to partially inactivate each gene in young adults (Supplemental Experimental Procedures). We found that depletion of FBF-1/2 strongly derepressed the gld-1 and fog-1 3′ UTR fusions but had no effect on the other fusions (Figure 5A and Table S2). Depletion of MEX-3 and GLD-1 caused all fusions to expand their expression distally. Most fusions maintained a small domain of no or low expression at the distal-most tip, except for rme-2, which was derepressed throughout the distal region (Figure 5A and Table S2). To confirm that expression in the distal-most region is due to derepression in germline progenitors and is not secondary to cell fate transformations, we repeated the RNAi depletions in L2 larvae, which contain only proliferating germ cells (Figure 5B and Table S3). These experiments confirmed that fog-1 and gld-1 require FBF-1/2 for repression in progenitors, whereas rme-2 requires GLD-1/MEX-3. None of the RNAi treatments affected the expression of the spe-11 promoter fusion in adults or in L2 larvae (Tables S2 and S3). We conclude that 3′ UTR-dependent regulation in progenitors involves inhibitory mechanisms mediated by FBF-1/2 and GLD-1/MEX-3.

FBF-1 and FBF-2 bind in vitro to a sequence motif (termed FBF-binding element or FBE) present in gld-1, fog-1, and lip-1 [17, 22]. Consistent with the RNAi results, we found that mutating the FBEs derepressed the gld-1 3′ UTR fusion but did not affect the lip-1 fusion (Figure S4). Unlike gld-1, which is repressed in a narrow distal domain contained within the mitotic zone, lip-1 is repressed in a broader region, perhaps reflecting more complex regulation (also see [23]). We note that FBEs were common in 3′ UTRs with no expression in the distal region (7/8), but were also present in a significant number of 3′ UTRs in our survey (17/30), including 3′ UTRs with strong expression in the distal region (5/8) (Table S4). Our results provide new evidence that, as suggested previously [18, 20], FBF-1/2 directly regulate the meiosis-promoting protein GLD-1 and that FBF-1/2 and MEX-3/GLD-1 regulate the expression of several genes via 3′ UTRs, although whether this regulation is direct remains to be tested.

Conclusions

In summary, our findings demonstrate that posttranscriptional mechanisms acting on 3′ UTRs are the main source of regulation throughout germ cell development, with the exception of...
spermatogenesis. FBF-1, FBF-2, GLD-1, and MEX-3 mediate 3' UTR-dependent repression in germline progenitors and are themselves dependent on 3' UTR sequences for proper expression (Figure 2), suggesting the existence of complex tiered networks of RNA-binding proteins and RNA target sites (see also [17]). Germ cells contain unique RNA-rich organelles (germ granules), which may represent a specialization for posttranscriptional regulation [24]. Whether similar RNA regulatory networks function in somatic lineages and how these interact with transcriptional networks [25] will be interesting to investigate. Changes in steady-state mRNA levels as detected by microarray analyses are often presumed to be due to transcriptional regulation but may also be due to posttranscriptional effects on mRNA stability [26]. Our survey demonstrates that 3' UTRs are a rich source of regulatory diversity that can bypass the need for most promoter specificity. Biochemical and computational methods have identified many potential functional sites in 3' UTRs, including binding sites for RNA-binding proteins and microRNAs [27–29]. Many such sites are present in the 3' UTRs in our survey (Table S4), although none correlate perfectly with a specific expression pattern. Systematic analysis of 3' UTR fusions, as initiated here, will be necessary to test the function of these sequences in vivo and elucidate the regulatory logic of 3' UTRs.

**Supplemental Data**

Supplemental Data include Supplemental Experimental Procedures, four figures, and six tables and can be found with this article online at [http://www.current-biology.com/cgi/content/full/18/19/1476/DC1/](http://www.current-biology.com/cgi/content/full/18/19/1476/DC1/).

**Acknowledgments**

We thank Jane Hubbard (NYU), Tim Schedl (Washington U.), Judith Kimble (U. of Wisconsin), and the anonymous reviewers for comments on the manuscript and Chris Gallo for pCG150. This work was supported by National Institutes of Health grant R01HD37047 to G.S., and National Institute of Child Health and Human Development training grant 2T32HD007276 to C.M. G.S. is an Investigator of the Howard Hughes Medical Institute.

Received: June 13, 2008
Revised: July 25, 2008
Accepted: August 5, 2008
Published online: September 25, 2008

**References**


