Correction

GENETICS

The authors note that an additional affiliation for Sebastian Hoersch was omitted from the article. Sebastian Hoersch’s second affiliation is Bioinformatics Group, Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany. The corrected author and affiliation lines appear below. The online version has been corrected.

Alla Grishok\textsuperscript{a}, Sebastian Hoersch\textsuperscript{a,b}, and Phillip A. Sharp\textsuperscript{a}

\textsuperscript{a}Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139; and \textsuperscript{b}Bioinformatics Group, Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany

www.pnas.org/cgi/doi/10.1073/pnas.1110993108
RNA interference and retinoblastoma-related genes are required for repression of endogenous siRNA targets in Caenorhabditis elegans

Alla Grishok,1,2, Sebastian Hoersch,1,2, and Phillip A. Sharp1,2

1Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139; and 2Bioinformatics Group, Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany

Contributed by Phillip A. Sharp, October 23, 2008 (sent for review August 15, 2008)

In Caenorhabditis elegans, a vast number of endogenous short RNAs corresponding to thousands of genes have been discovered recently. This finding suggests that these short interfering RNAs (siRNAs) may contribute to regulation of many developmental and other signaling pathways in addition to silencing viruses and transposons. Here, we present a microarray analysis of gene expression in RNA interference (RNAi)-related mutants rde-4, zfp-1, and alg-1 and the retinoblastoma (Rb) mutant lin-35. We found that a component of Dicer complex RDE-4 and a chromatin-related zinc finger protein ZFP-1, not implicated in endogenous RNAi, regulate overlapping sets of genes. Notably, genes a) up-regulated in the rde-4 and zfp-1 mutants and b) up-regulated in the lin-35(Rb) mutant, but not the down-regulated genes are highly represented in the set of genes with corresponding endogenous siRNAs (endosRNAs). Our study suggests that endogenous siRNAs cooperate with chromatin factors, either C. elegans ortholog of acute lymphoblastic leukemia-1 (ALL-1)-fused gene from chromosome 10 (AF10), ZFP-1, or tumor suppressor Rb, to regulate overlapping sets of genes and predicts a large role for RNAi-based chromatin silencing in control of gene expression in C. elegans.

endo-siRNA  | microarray  | Rb  | RNAi

Among the species with sequenced genomes the nematode Caenorhabditis elegans encodes the largest number of Argonaute proteins, which interact with short RNAs (1). Also, a large number of endogenous, short interfering RNAs (endo-siRNAs) have been cloned from C. elegans (2–5). They are distinct from miRNAs, are largely generated by RNA-dependent RNA polymerases (RdRP), and match thousands of genes. These observations suggest that multiple gene-regulatory networks involving Argonaute proteins and endosRNAs exist in the nematode.

We have characterized a system of RNAi-induced transcriptional gene silencing (RNAi-TGS) of a repetitive transgene expressed in the soma of C. elegans (6). Also, we found that RNAi pathway genes and lin-35(Rb) synergize in repressing the intestinal cell divisions and in repressing the cyclin E gene (cye-1) expression, likely through cooperative inhibition of cye-1 transcription (7). Two chromatin-related genes, zfp-1 and gfl-1, promote the RNAi process in C. elegans, either directly or indirectly, they also contribute to RNAi-TGS of a repetitive transgene (6, 8, 9). Interestingly, both genes were also found to antagonize the repressive function of LIN-35(Rb) (10, 11). Therefore, ZFP-1 and GFL-1 appear to regulate both RNAi and Rb target genes.

The C. elegans Rb protein LIN-35 represses inappropriate transcription of germline-specific genes (12) and growth factors (13) in differentiated somatic cells and functions redundantly with other transcriptional repressors (14). Also, lin-35 mutants are more sensitive to exogenous RNAi than wild-type worms (11, 15). This might be partially because of the de-repression of germline-specific RNAi pathway genes in somatic cells.

Because RNAi genes were found to function in the same processes as lin-35, we conducted microarray experiments to find potential targets regulated by RNAi-TGS and lin-35. We used rde-4 and zfp-1 mutants affecting RNAi-TGS. RDE-4 is a dsRNA binding protein interacting with Dicer (16) whereas ZFP-1 is a nuclear protein that is likely to affect transcription directly. Our previous study indicated that miRNAs might have a role in promoting RNAi-TGS in C. elegans as well (6); therefore, we included miRNA pathway Argonaute mutant alg-1 in our experiments.

Our analysis revealed i) that zfp-1 and rde-4 mutant animals have strikingly similar profiles of alterations in gene expression and ii) that there is an enrichment of genes with matching (antisense) endosRNAs (3–5) only among genes up-regulated, but not down-regulated, in zfp-1 and rde-4 mutants. These genes therefore might represent direct targets of chromatin-based silencing induced by endogenous RNAi pathways. Interestingly, endosRNAs matched not only genes negatively regulated by rde-4 and zfp-1, but also those primarily inhibited by LIN-35(Rb).

We also report that zfp-1, unlike rde-4, opposes the repressive function of LIN-35 in controlling intestinal nuclear divisions and cye-1 expression. Our results suggest that ZFP-1 may play both a positive and a negative role in regulating gene expression.

Results

Microarray Data Analysis. To find target genes regulated by RNAi and Rb, we performed a series of microarray experiments using RNA from L1-L2 larvae of the wild type and loss-of-function mutants rde-4 (17), zfp-1 (10), alg-1 (7), and lin-35 (18). We conducted pairwise comparisons of the levels of gene expression in each mutant compared with the wild type and selected statistically significant changes in gene expression by two-sample t test (P value <0.01), requiring in addition an expression difference of at least 1.5-fold between two group averages. Our microarray data are summarized in Dataset S1 and Dataset S2.

A majority of the genes changing expression in the lin-35 mutant compared with the wild type (535 of 710) were up-regulated consistent with the repressive role of the LIN-35 protein (Table 1). Similar numbers of genes were either up-regulated or down-regulated in each of the RNAi-related mutants: 420 were “up” in zfp-1 and 434 were “down” whereas 285 were “up” in rde-4 and 219 were “down”, and 170 were “up” in alg-1 and 213 were “down.” The numbers of genes similarly regulated in different mutants are listed in Table 1. Ten genes commonly up-regulated in all four mutants are described in Table S1.

Acknowledgments

Author contributions: A.G., S.H., and P.A.S. designed research; A.G. performed research; A.G. and S.H. analyzed data; and A.G., S.H., and P.A.S. wrote the paper. The authors declare no conflict of interest.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo accession no. GSE13258.

1Present address: Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, New York, NY 10032.
2To whom correspondence should be addressed at: Koch Institute for Integrative Cancer Research, MIT, 40 Ames Street, E17-529, Cambridge, MA 02139. E-mail: sharppa@mit.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0810589105/DCSupplemental.

© 2008 by The National Academy of Sciences of the USA
related datasets are clustered on the mutants and close to 20% of genes commonly up-regulated in both mutants in specific groups are listed in Dataset S3.

Statistical significance and representation factors for gene enrichment and comparison of multiple datasets. P-values for categories to 5,615 functionally characterized C. elegans genes (77% for topomap compared with 46% for GO) that is—by nature of the “annotation process”—not restricted to known and characterized genes. Therefore, topomap-based functional assignment described in our study is not limited to well studied genes. Functional annotation of our expression data using GO platform (data not shown) was similar to that obtained with topomap, but we arrived at a more complete picture of gene expression by using topomap.

A comparison of the functional categories of genes changing expression in different mutants revealed a striking similarity between transcriptome profiles in rde-4 and zfp-1 mutants (Fig. 1 and Fig. S1). This similarity suggests that common biological processes are affected by both mutations. For example, certain germline-enriched and oocyte genes (mount #02) are overrepresented in groups of genes with higher expression levels in zfp-1 and rde-4 mutants and close to 20% of genes commonly up-regulated in both mutants belong to this category (Dataset S2). Indeed, functional annotation of the groups of genes commonly affected by each combination of two mutants (presented in Table 1) revealed the same categories of enrichment as those that were common between the two single mutant profiles (Fig. S2).

Therefore, two independent types of analyses: 1) a direct comparison of genes changing expression in two mutants (Table 1) and 2) functional annotation of misregulated genes (Fig. 1 and Figs. S1 and S2) strongly suggest that zfp-1 and rde-4 work in the same pathway (RNAi-TGS) and point to a very significant role of this pathway in biology of C. elegans.

The rde-4 mRNA level was not changed in the zfp-1 mutant and vice versa, indicating that a simple model of regulation of one gene by the product of another does not account for the correlation. We cannot exclude the possibility that protein levels of RDE-4 or ZFP-1 might change. However, these types of changes are not likely to be due to the direct regulation by RDE-4 or ZFP-1 because RDE-4 is known to interact with RNA and ZFP-1 is a chromatin factor.

Genes with higher expression in zfp-1 and rde-4 mutants were overrepresented among the functional groups ‘protein expression,’ ‘germline-enriched,’ ‘biosynthesis,’ ‘mitochondrial,’ and ‘cell cycle,’ whereas those genes that were down-regulated in the mutants frequently represented intestine-specific genes involved in metabolic processes (Fig. 1). Histone genes were also significantly enriched in the rde-4 down-regulated gene set (Fig. 1). Importantly, ZFP-1 appears to have a larger role in gene expression regulation than RDE-4 (Fig. 1, Table 1, and Dataset S1). Consistent with these results, zfp-1 mutants have some developmental phenotypes, such as slow growth and protruded vulva (10), whereas rde-4 mutant worms are superficially normal.

A recent microarray study reported gene expression changes in the RNAi pathway mutants rde-1, rde-4, and dcr-1 (20). We mapped the misregulated gene sets from this study to the functional groups of coregulated genes (Fig. 1) and found that genes down-regulated in the rde-4 mutant were enriched in intestine-specific group contained significant number of histone genes and proteases. This signature corresponds to that of genes down-regulated in the rde-4 mutant from our study (Fig. 1). However, genes found up-regulated in the rde-4 mutant do not have a signature consistent with our findings (Fig. 1). One difference between the studies is that we used L1-L2 larva and the published report used adult worms (20). Because adult worms contain both differentiated somatic tissues and actively proliferating and specialized germine cells, whereas the L1-L2 larvae contain primarily somatic cells, the resulting “average” gene expression profile is likely to be different in adults and larvae. In addition, mutant backgrounds may have different effects on gene expression in somatic and germine tissues.

Table 1. Numbers of genes changing expression compared with the wild type in indicated mutant backgrounds (top) and numbers of overlapping genes between indicated mutants (bottom)

<table>
<thead>
<tr>
<th>Mutant backgrounds</th>
<th># genes UP</th>
<th>Enrichment</th>
<th>P value</th>
<th># genes DOWN</th>
<th>Enrichment</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>lin-35</td>
<td>535</td>
<td></td>
<td></td>
<td>175</td>
<td></td>
<td></td>
</tr>
<tr>
<td>zfp-1</td>
<td>420</td>
<td></td>
<td></td>
<td>434</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rde-4</td>
<td>285</td>
<td></td>
<td></td>
<td>219</td>
<td></td>
<td></td>
</tr>
<tr>
<td>alg-1</td>
<td>170</td>
<td></td>
<td></td>
<td>213</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lin-35 and zfp-1</td>
<td>56</td>
<td>4.54</td>
<td>1.39 x 10^-21</td>
<td>65</td>
<td>15.59</td>
<td>2.15 x 10^-60</td>
</tr>
<tr>
<td>lin-35 and rde-4</td>
<td>40</td>
<td>4.78</td>
<td>2.07 x 10^-16</td>
<td>32</td>
<td>15.21</td>
<td>9.1 x 10^-29</td>
</tr>
<tr>
<td>lin-35 and alg-1</td>
<td>39</td>
<td>7.81</td>
<td>5.45 x 10^-24</td>
<td>40</td>
<td>19.55</td>
<td>1.58 x 10^-40</td>
</tr>
<tr>
<td>zfp-1 and rde-4</td>
<td>110</td>
<td>16.74</td>
<td>1.44 x 10^-107</td>
<td>138</td>
<td>26.45</td>
<td>2.78 x 10^-174</td>
</tr>
<tr>
<td>zfp-1 and alg-1</td>
<td>68</td>
<td>17.35</td>
<td>7.26 x 10^-67</td>
<td>131</td>
<td>25.81</td>
<td>7.40 x 10^-163</td>
</tr>
<tr>
<td>rde-4 and alg-1</td>
<td>41</td>
<td>15.41</td>
<td>3.77 x 10^-37</td>
<td>77</td>
<td>30.07</td>
<td>2.65 x 10^-96</td>
</tr>
</tbody>
</table>

Genes up-regulated in the mutants are listed in the ‘UP’ column and down-regulated gene are listed in the ‘DOWN’ column. Enrichment factors and P-values for significance of overlaps are indicated. Descriptions of genes changing expression in indicated mutants are listed in Dataset S2. Overlaps between groups of genes misregulated in two or more mutants can be found by using the “data filter” option in the Excel file and selecting ‘1’ in the columns representing conditions of interest.
was nonrandom and mostly consistent among the three studies that mapped the endo-siRNA target genes according to three independent studies (3–5) (Fig. 1). We found that siRNA-matching genes were very significantly enriched in the same functional groups as the genes up-regulated in *zfp-1* and *rde-4* mutants or *lin-35* mutants (Fig. 1), but not in the groups overrepresented in genes down-regulated in the same mutants, with the exception of histone genes (Fig. 1). This suggested that specific genes negatively regulated by *zfp-1* and *rde-4* or by *lin-35* might be more likely to have a matching endo-siRNA.

Indeed, direct comparison of endo-siRNA target gene sets and *zfp-1*, *rde-4*, *alg-1*, and *lin-35*-regulated genes revealed a statistically significant enrichment of genes with siRNAs only in the sets of genes up-regulated in the mutants but not in the down-regulated sets of genes (Table 2). Close to 50% of genes with increased expression in each of the mutant strains were reported to have a matching siRNA. These data strongly suggest that genes up-regulated in the mutants represent the direct targets repressed by RNAi and that the down-regulated genes might be affected by the mutations indirectly. A very large overlap between *rde-4* and *zfp-1*-regulated genes (Table 1) together with high significance of enrichment in siRNA targets of gene sets up-regulated in both mutants (Table 2) further strengthens the prediction of numerous target genes regulated by RNAi-TGS in *C. elegans*.

When functional annotation was done on the groups of genes representing overlaps between genes up-regulated in the mutants and siRNA target genes, the signatures of “UP in *zfp-1* and siRNA target” and “UP in *rde-4* and siRNA target” groups were found to be very similar and very close to “UP in *zfp-1* and *rde-4*” signature. On the contrary, the signature of the “UP in *lin-35* and siRNA target” group was distinct from “UP in *zfp-1* and *rde-4*” and very similar to that of “UP in *lin-35*”, whereas “UP in *alg-1* and siRNA target” group had similarity to groups of genes regulated by *zfp-1* and *rde-4* and also regulated by *lin-35*. We conclude that siRNA targets overlapping with *lin-35*-regulated genes are distinct from the groups of siRNA targets regulated by *zfp-1* and *rde-4*. Although chromatin factor ZFP-1 may be directly involved in the endo-siRNA pathway as this gene was implicated in supporting RNAi, the overlap between *lin-35* and endo-siRNA target genes likely represents synergy between the two repressive pathways.

The main signature of the *lin-35* mutant is de-repression of germ-line-specific genes in somatic tissues of larvae (Fig. 1), which is consistent with previous findings (11, 22). There are three main groups of coregulated genes that represent germ line (mounts #02, #07, and #11, Fig. S1). Endogenous siRNAs are enriched in those same groups: of the total of 4,372 siRNA targets represented in the topomap dataset, 1,448 were found to belong to these germ-line-specific genes in somatic tissues of larvae (154 compared with 17 expected by chance). This correlation may indicate that endo-siRNAs synergize with LIN-35 in repressing germ-line-specific fate in somatic tissues. Alternatively, although both LIN-35 targets and endo-siRNAs preferentially correspond to germ-line-enriched genes, LIN-35 and RNAi may regulate those genes independently in distinct tissues: soma and germ line, respectively.

When we compared nongermine siRNA target genes with gene sets changing expression in the mutants, genes up-regulated in *lin-35* were enriched modestly (2.5×) and enriched less than genes down-regulated in *lin-35* (3.3×) (Table S3). In contrast, genes up-regulated in *zfp-1* and *rde-4* were overrepresented among siRNA targets (4×), independently of their germine or nongermine classification (Tables S2 and S3). The corresponding sets of down-regulated genes were not overrepresented. This analysis further supports synergy between endo-siRNAs, *rde-4* and *zfp-1*, in gene expression regulation. Although *lin-35* and endo-siRNA targets do not appear to correlate outside of germ-line-enriched group of genes, the possibility of synergy between LIN-35 and endo-siRNAs in repressing germine fate in the soma still remains and needs to be studied further.

---

**Fig. 1.** Heatmap showing the enrichment of selected functional groups as defined by ref. 19 (y axis) in gene sets from various datasets (x axis). The gray shades in the heatmap indicate significance levels as indicated in the legend. The bars to the right of the heatmap indicate the total size of each functional group. For a comprehensive version of the figure including all functional groups with any significant enrichment and clustering dendrograms informing the order of groups and datasets, see Fig. S1. Enrichment factors and their P values are presented in Dataset S3.

**Endogenous siRNA Preferentially Map to Genes Up-Regulated in zfp-1, rde-4, and lin-35 Mutants.** As mentioned, we found almost equal numbers of genes both positively and negatively regulated by RDE-4 and ZFP-1 (Table 1). Although it is conceivable that ZFP-1 might act directly as an activator or as a repressor on both types of target genes, RDE-4, required for the production of siRNAs (21), is much more likely to contribute directly only to gene silencing.

To gain insight about possible direct targets of RNAi-TGS, we mapped the endo-siRNA target genes according to three independent studies (3–5) to the coregulated groups defined by Kim and colleagues (Fig. 1). Interestingly, distribution of siRNA target genes was nonrandom and mostly consistent among the three studies (Fig. 1). When functional annotation was done on the groups of genes representing overlaps between genes up-regulated in the mutants and siRNA target genes, the signatures of “UP in *zfp-1* and siRNA target” and “UP in *rde-4* and siRNA target” groups were found to be very similar and very close to “UP in *zfp-1* and *rde-4*” signature. On the contrary, the signature of the “UP in *lin-35* and siRNA target” group was distinct from “UP in *zfp-1* and *rde-4*” and very similar to that of “UP in *lin-35*”, whereas “UP in *alg-1* and siRNA target” group had similarity to groups of genes regulated by *zfp-1* and *rde-4* and also regulated by *lin-35*. We conclude that siRNA targets overlapping with *lin-35*-regulated genes are distinct from the groups of siRNA targets regulated by *zfp-1* and *rde-4*. Although chromatin factor ZFP-1 may be directly involved in the endo-siRNA pathway as this gene was implicated in supporting RNAi, the overlap between *lin-35* and endo-siRNA target genes likely represents synergy between the two repressive pathways.

The main signature of the *lin-35* mutant is de-repression of germ-line-specific genes in somatic tissues of larvae (Fig. 1), which is consistent with previous findings (11, 22). There are three main groups of coregulated genes that represent germ line (mounts #02, #07, and #11, Fig. S1). Endogenous siRNAs are enriched in those same groups: of the total of 4,372 siRNA targets represented in the topomap dataset, 1,448 were found to belong to these germ-line-specific genes in somatic tissues of larvae (154 compared with 17 expected by chance). This correlation may indicate that endo-siRNAs synergize with LIN-35 in repressing germ-line-specific fate in somatic tissues. Alternatively, although both LIN-35 targets and endo-siRNAs preferentially correspond to germ-line-enriched genes, LIN-35 and RNAi may regulate those genes independently in distinct tissues: soma and germ line, respectively.

When we compared nongermine siRNA target genes with gene sets changing expression in the mutants, genes up-regulated in *lin-35* were enriched modestly (2.5×) and enriched less than genes down-regulated in *lin-35* (3.3×) (Table S3). In contrast, genes up-regulated in *zfp-1* and *rde-4* were overrepresented among siRNA targets (4×), independently of their germine or nongermine classification (Tables S2 and S3). The corresponding sets of down-regulated genes were not overrepresented. This analysis further supports synergy between endo-siRNAs, *rde-4* and *zfp-1*, in gene expression regulation. Although *lin-35* and endo-siRNA targets do not appear to correlate outside of germ-line-enriched group of genes, the possibility of synergy between LIN-35 and endo-siRNAs in repressing germine fate in the soma still remains and needs to be studied further.
Notably, the cyclin E gene targeted by endo-siRNAs is expressed very highly during oogenesis and is categorized as ‘germline-enriched.’ Therefore, its repression by Rb and RNAi pathways in somatic tissues that we discovered genetically (7) may serve as an example of possible large-scale cooperation between endo-siRNAs and LIN-35 in repressing common targets.

**Mutation in zfp-1 Suppresses Extra Nuclei Division Phenotype and Enhanced Cyclin E Expression in lin-35 Mutant Worms.** We have found that the combination of the RNAi pathway mutants rde-1, rde-4, or the miRNA pathway mutants dcr-1 and alg-1 with the lin-35 mutation leads to a significant increase in postembryonic nuclear divisions in the intestine of the double mutant worms (Fig. 2A and B) (7). Increases in cyclin E (cye-1) transcription under these conditions are at least partially responsible for this phenotype.

Because rde-4 and zfp-1 regulate many common genes, we tested whether ZFP-1 also cooperates with LIN-35 in repressing cye-1. Surprisingly, we found that combining the zfp-1 mutation with lin-35(lf) did not lead to an increase in nuclear divisions. Instead, the zfp-1 mutation suppressed extra nuclear divisions associated with the lin-35; dcr-1, lin-35; alg-1 (Fig. 2A) and lin-35; rde-1 double mutant combinations (Fig. 2B). This suppression by zfp-1(lf) of a phenotype associated with the lack of transcriptional repressors is comparable with its suppression of a multivulva phenotype (10, 11). In both cases it is likely that zfp-1 function is required for an enhanced expression of the de-repressed target genes.

Because cyclin E is one of the target genes repressed by LIN-35, we tested whether enhanced expression of cye-1 mRNA in lin-35(lf) worms requires ZFP-1. Indeed, we found that in the lin-35; zfp-1 double mutant strain, the cye-1 mRNA level was reduced as compared with that in lin-35(lf) (Fig. 2C). We did not observe a reduction in cye-1 mRNA levels in the zfp-1 mutant alone, indicating that its activity is not required for normal levels of expression of this gene.

**Genes Repressed by lin-35 and Activated by zfp-1.** Our genetic studies of cye-1 regulation and published reports (10, 11) indicate that zfp-1 may act as an activator of LIN-35(Rb)-repressed genes (Fig. 2). However, the microarray results strongly suggest that ZFP-1 and RDE-4 have a direct repressive effect on a number of other targets, which are not regulated by LIN-35. We were interested in identifying an additional group of genes, those oppositely regulated by lin-35 and zfp-1, and further selected for up-regulated expression in lin-35(lf) background and down-regulated expression in the zfp-1 mutant with a change in expression intermediate between zfp-1 and lin-35 in rde-4(lf) and alg-1(lf) (see SI Text). Fifty-seven genes with expression profiles showing high similarity to this “custom expression profile” were identified (Dataset S1 and Fig. 1). Notably, three Argonaute genes were found in this group. This representation is statistically significant (enrichment factor 38x, P value 6.36 × 10^{-5}).

We used quantitative real-time PCR to analyze the expression levels of the candidate genes with the largest differences in expression between zfp-1 and lin-35 mutants (down-regulated in zfp-1 and up-regulated in lin-35) or genes with smaller expression changes that we find interesting, such as Argonaute gene ccr-1 (1). The expression of these chosen genes was tested in mutants used for the array analysis and in lin-35; rde-4 and lin-35; zfp-1 double mutants that have limited viability (Fig. 3 A–F). A few genes showed suppression of their enhanced expression in lin-35(lf) background.

---

**Table 2. endo-siRNA targets are overrepresented among genes up-regulated in RNAi and Rb mutants**

<table>
<thead>
<tr>
<th>UP in lin-35</th>
<th>UP in zfp-1</th>
<th>UP in rde-4</th>
<th>UP in alg-1</th>
<th>DOWN in lin-35</th>
<th>DOWN in zfp-1</th>
<th>DOWN in rde-4</th>
<th>DOWN in alg-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>428</td>
<td>333</td>
<td>232</td>
<td>142</td>
<td>143</td>
<td>371</td>
<td>191</td>
<td>185</td>
</tr>
</tbody>
</table>

Endo-siRNA target gene set, according to Pak and Fire (2) and gene sets determined to be “UP”- or “DOWN”-regulated in our expression data were mapped to 18,459 genes with TOPOMAP representation (with recalls ranging from 74% to 84%). For each group, total and overlap counts are listed as well as representation factors and p values for overlaps.
when the zfp-1 mutation was added to the lin-35 mutant (Fig. 3 E and F), whereas enhanced expression of other genes was not suppressed by zfp-1( lfl) (Fig. 3 A–D). These results reveal a complex regulation of tested genes by LIN-35 and ZFP-1 and suggest that ZFP-1 may have a dual role (of an activator and repressor) in regulating expression of specific targets.

We also performed real time RT-PCR analysis of the expression of several germline-enriched genes with matching siRNAs and repressed by lin-35, but not affected by zfp-1(o5554), similarly to cye-1 (Fig. S3). The lin-35; rde-4 and lin-35; zfp-1 double mutants were also included in this analysis. Loss of rde-4 or zfp-1 does not appear to contribute very significantly to the dramatic de-repression of these targets in the lin-35 mutant background.

Discussion

Our microarray study was motivated by the finding of cooperation between RNAi-TGS and Rb in cyclin E regulation (7). We aimed at identifying more targets of these repressive pathways.

The profile of genes up-regulated in the lin-35 mutant larva confirms its role in the repression of germline-specific fates in somatic cells (11). More than half of genes up-regulated in the lin-35 mutant have matching endogenous siRNAs (311/535, enrichment factor 2.3×, P value 2 × 10^{-6}). Although we cannot exclude a possibility that these endo-siRNAs are produced in the germ line and also function in this tissue, it is equally possible that siRNAs generated in the germ line by RdRPs are inherited and function along with LIN-35 to repress germline-specific genes in the soma. Cyclin E is an example of a ‘germline-enriched’ gene repressed by LIN-35 and RNAi in the somatic tissues (7). This pattern of expression of Cyclin E is not unique to nematodes. Cyclin E expression has been shown to be continuous throughout the cell cycle in germline stem cells of Drosophila (23) and embryonic stem (ES) cells from mouse (24) and primates (25). High level of cyclin E was proposed to indicate “stemness” of the cells (23). In somatic cells in these organisms, constitutive Cyclin E expression is repressed with the onset of cell-cycle dependent regulation. Our results demonstrating repression of cyclin E in the soma along with other germine genes are consistent with this idea.

Interestingly, we found that many RNAi-related Argonaute genes (ppw-1, sago-2, C16C10.3, C04F12.1, and csr-1) are repressed by LIN-35. Argonaute proteins interact with siRNAs and are essential for the silencing process. C. elegans Argonaute genes ppw-1, sago-2, and C04F12.1 function redundantly in the RNAi process (1). The level of expression of these genes is elevated eight to ten fold in lin-35(lf) larvae. This finding may explain why the lin-35 mutant is more susceptible to exogenous RNAi (11, 15).

We identified very significant enrichment of endo-siRNA target genes among genes up-regulated in rde-4 (P value 9.6 × 10^{-7}) and zfp-1 (P value 1.4 × 10^{-2}) mutants. Also, these mutants affected a large number of common genes. Previous studies aimed at identifying common misregulated targets among various endo-RNAi pathway mutants (3) have not detected large overlaps in misregulated genes or common functional signatures predicting biological pathways where regulation by endogenous RNAi may take place. Therefore, this is the first study demonstrating a connection between zfp-1 function and endogenous RNAi processes and identifying specific genes that are 1) endo-siRNA targets, 2) up-regulated in rde-4(−/−), and 3) up-regulated in zfp-1(−/−) and belong to very specific functional groups, such as regulation of protein translation and germline function (Dataset S2).

We infer that genes commonly up-regulated in the rde-4 and zfp-1 mutants and containing matching siRNAs are the direct targets of nuclear RNAi. This prediction is based on the role of rde-4 and zfp-1 genes in our characterized system of transcriptional silencing of a transgene (6), the demonstrated requirement of RDE-4 for production of at least some endo-siRNAs (3, 26) and on the predicted nuclear function of the ZFP-1 protein. ZFP-1 is a homolog of mammalian protein AF10, which causes myeloid leukemia when fused to MLL (27). Both ZFP-1 and AF10 contain two N-terminal PHD zinc fingers and a C-terminal leucine zipper domain. Some PHD zinc fingers were recently recognized as histone-binding modules interacting with either methylated (28, 29) or unmethylated (30) lysine 4 of histone H3. The protein sequences of most terminal PHD fingers of ZFP-1 and AF10 align very well with histone-binding PHD fingers of other proteins, strongly suggesting that these proteins interact with chromatin via PHD domains. AF10 was shown to recruit histone H3 lysine 79 Dot1 methyltransferase via its leucine zipper domain (31) and to play a role in transcriptional elongation (32). It is possible that ZFP-1 binds histones with its N-terminal PHD domain and recruits different protein factors with its C-terminal domain. It could serve as an adaptor for both activators (Dot1) and repressors (RNAi factors) and regulate gene expression at the transcription elongation step.

The majority of the endo-siRNAs in C. elegans is antisense to mature mRNA sequences and is likely produced by RdRPs by using those mature RNAs as templates (2–5). A very recent discovery of an Argonaute protein NRDE-3 that binds endo-siRNAs and shuttles between the cytoplasm and the nucleus (26) further supports a
and used instruments to facilitate gene expression comparison between a RNA Preparation and Microarray Hybridization.

Materials and Methods

C. elegans Strains. Worms were maintained on nematode growth medium plates seeded with OP50 bacteria. The strains used are listed in the SI Text. Adult or L4 worms were used for counting intestinal nuclei in strains containing elt-2:gfplacZ reporter. RNAi by feeding was performed as described (7). We used lin-35(n745) mutant linked to the weak unc-13(e1091) allele in our experiments to facilitate gene expression comparison between a lin-35 single mutant and lin-35; rde-4 and lin-35; zfp-1 double mutants constructed in unc-13(e1091) background. Only one of eighteen lin-35unc-13-dependent genes that we tested by real-time RT-PCR, sod-3, had an increased expression in unc-13(e1091) background compared with wild type (data not shown). However, its expression was even higher in the lin-35unc-13 strain. Because the functional categories of genes up-regulated in lin-35 mutant were almost identical between our study and that of Kirienko and Fay (22) (Fig. 1), which used an unmarked lin-35 mutant, we believe that the number of false positives in our study, due to unc-13, is very low.

C. elegans Collection for Microarray Experiments. Nematodes were synchronized at L1 stage by hypochlorite treatment of gravid hermaphrodites and hatching their eggs overnight in liquid culture without food. Resulting pop-

RNA Preparation and Microarray Hybridization. Tri Reagent (MRC) was used for total RNA preparation from frozen worms resulting in 5–30 μg RNA per sample. The quality of RNA samples was confirmed by BioRad Bioanalyzer. Affymetrix GeneChip C. elegans Genome Arrays with a total of 22,625 probesets were hybridized with cDNA and scanned according to manufacturer’s standard protocol. All conditions (WT and 4 mutants) were profiled in triplicate. Replicates were biological replicates (separately grown worm populations), with two exceptions: because of shortage of biological material, there were only two bio-

Data Analysis: Sets of Differentially Expressed Genes. Sets of probesets with up-
or down-regulated expression in the mutants relative to WT were determined via t test (two-tailed, homoscedastic) with a P value cutoff of 0.01, requiring in addition an average expression difference of 1.5 or greater on the natural scale. Complete data analysis description, which includes generation of idealized expression profile, gene assignment and mapping, topomap assignments and graphic generation, is presented in SI Text.

RT and quantitative real time PCR was performed as described in refs. 6 and 7.

ACKNOWLEDGMENTS. We thank Manlin Luo for performing cDNA labeling and microarray hybridizations and Charlie Whittaker for help with microarray data processing. We also thank Iva Greenwald, Oliver Hobert, Joel Neilson, and Anthony Leung for comments on the manuscript. The elt-2:gfplacZ strain was generated by Anne Hart (Massachusetts General Hospital). The zfp-1(ok554) strain was provided by the C. elegans Gene Knockout Project at Oklahoma Medical Research Foundation, which is part of the International C. elegans Gene Knockout Consortium. Some strains used in this study were obtained from the C. elegans Genetics Center, which is funded by National Institutes of Health National Center for Research Resources. This work was supported by a Leukemia and Lymphoma Foundation Fellowship #3260–07 (to A.G.), United States Public Health Service Grant PO1-CA42063 from the National Cancer Institute (to P.A.S.), and partially by Cancer Center Support Grant P30-CA14501 from the National Cancer Institute.


