Supporting Online Material for

Distinct Populations of Primary and Secondary Effectors During RNAi in *C. elegans*

Julia Pak and Andrew Fire*

*To whom correspondence should be addressed. E-mail: afire@stanford.edu

Published 23 November 2006 on *Science* Express
DOI: 10.1126/science.1132839

This PDF file includes:

Materials and Methods
SOM Text
Figs. S2 to S6
Tables S1 and S2
References

Other Supporting Online Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/1132839/DC1)

Fig. S1 as zipped archive 1132839s1.zip
Materials and Methods

Nematode growth and harvesting
Wild type N2 animals were grown on large (9cm) NGM plates containing 25 µg/ml carbenicillin, 232.5 µg/ml IPTG and HT115 bacteria containing the pJP12.2+ plasmid (containing the sel-1 trigger sequence) as a food source. pJP12.2+ is a derivative of the C. elegans dsRNA feeding vector L4440 (1) carrying a 458 bp segment of the sel-1 coding sequence (TGCA TGACGAAGTTTGAGCAGA to AGTCAGGAAGTTTGAGCAGA). sel-1 was chosen as an RNAi trigger based on the widespread expression of the endogenous target gene. During feeding-based RNAi, the bulk of sel-1 siRNAs appear to result from soma-specific RdRP activity (the rrf-1 gene product), as evidenced by a substantial decrease in the siRNA signal (on Northern blots) in an rrf-1 null mutant background and the lack of observed alteration in signal for animals lacking the majority of germline tissue (glp-4 (bn2) background) (2, 3, data not shown). The effectiveness of the RNAi trigger was confirmed by functional suppression of mutations in the interacting gene glp-1 (4, 5). Following 7-10 days at 16ºC, C. elegans populations were washed free of bacteria in 5% sucrose, 50mM NaCl, 5mM NaEDTA (pH 8) and frozen as pellets in liquid nitrogen. RNA was extracted by crushing the frozen samples using a mortar and pestle before extraction using the miRvana™ miRNA Isolation Kit (Ambion).

Northern blot analysis
Three gel systems were used that differ in acrylamide:bis-acrylamide ratio; 20% 19:1 acrylamide:bis-acrylamide, 20% 75:1 acrylamide:bis-acrylamide, and 12% 9:1 acrylamide:bis-acrylamide. In all cases, the gel contained 8M urea and 100mM Tris, 90mM Boric Acid, 1mM EDTA (1 x TBE) except for the 75:1 acrylamide:bis-acrylamide gel which was run in 50mM Tris, 45mM Boric Acid, 0.5mM EDTA. We have found that under conditions of a high acrylamide:bis-acrylamide ratio (as in the 75:1 gel), gel migration of RNA can be paradoxically affected by 5’-modification. On 75:1 gels, 5’-monophosphate and 5’-triphosphate results in similar gel mobilities, both migrating distinctly faster than the corresponding 5’-OH RNAs (Fig. 3C). Using an acrylamide:bis-acrylamide ratio of 9:1 (Fig. 3D), gel mobilities more directly resolve small RNAs based on the numbers of 5’-phosphates.

sel-1 probes for Northern blots were 60nt oligonucleotides corresponding to the sel-1 sequence contained in the trigger plasmid (pJP12.2+) overlapping by 20nt that were end-labeled with T4 polynucleotide kinase and gamma-32P-ATP;

5’ GACTAACCAGAACCAGGATCCAATCCATGCTAACCAGGAAGTGACCACTGGAGA 3’
5’ AGGAAGTGACCACTGGAGAGATATTTCTTGAGTATTACAGATGTGTTCGGACAAAGG 3’
5’ AAGATGCTTGGCCGACAAGAAGGAGATAACATGCCCTCAATTGGGACTCGGACAGAGTTATTTA 3’
5’ GACTCGGACAGATTATTTACGCTGTCGAGGAGGCTCAATCCAATTTCTCAACTCAGGACTCGCCT 3’
5’ TCACAAATTTSCGAAACTCCCGCTTCCCGACTTTTCTGCTGTAGTCTAGGAGCTGCCGAAGCTGGCC 3’

The Northern blot shown in Fig. 3D was probed with an additional 32P-end labeled oligonucleotide (5’ TGTAGTGAAATCTTTAAATCC 3’) that corresponds to the sel-1 sequence comprising the in vitro transcription products.
Enzymatic assays of small RNAs

For enzymatic analysis of structures, small RNA preparations were treated with a variety of enzymes under manufacturer-recommended conditions. Where indicated in the figures, 10µg of size-selected small RNAs were treated with calf intestinal alkaline phosphatase (CIP, Amersham), T4 polynucleotide kinase (T4 PNK, New England Biolabs), and/or T4 RNA Ligase (New England Biolabs) at 37°C for one hour. Ligated products (from natural or synthetic RNA) are most likely heterogeneous in nature and may be limited in their filter-binding ability. These RNAs were poorly detected, with their presence indicated by the loss of the unligated species. RNase T (New England Biolabs) was added to samples for one hour at room temperature, and Terminator (Epicentre) was added for one hour at 30°C. Where indicated, internal 25nt synthetic sel-1 oligonucleotide controls were added to enzymatic reactions of in vivo small RNAs.

In vitro transcription reactions

RNA was in vitro transcribed using T7 RNA polymerase (Ambion) from a short double-stranded DNA template comprising the T7 promoter and 22 nt of sel-1 sequence that has no internal guanosine residues; 5’GGTAATACGACTCACTATAGAATTTAACATTATCCACTACA3’ and 5’TGTAGTGGATAATGTTAAATTCTATAGTGAGTCGTATTACC3’.

The transcription reaction took place in the presence of ATP, UTP, CTP and GTP or with GTP replaced with guanosine monophosphate, guanosine diphosphate, or guanosine tetraphosphate. The 5’-OH product was generated by alkaline phosphatase treatment of the 5’-triphosphate product. The transcription reactions resulted in a number of minor bands that were produced as well as the major species. These may reflect imprecisely initiated or terminated transcripts, nucleolytic degradation, or heterogeneity in the DNA oligonucleotide templates used in the reactions.

5’-ligation-independent cloning

The following protocol is a modification of that designed by Lau et al. (6). 200µg of small RNAs derived from the miRVana isolation kit were first further size-selected on a 12% denaturing polyacrylamide gel (size range: approximate 18-25nt). The extracted product was ligated to a 3’ linker that has an adenylated 5’-end and a ddC at the 3’-end as described in Lau et al. (6). This was followed by a second gel purification. The ligated product was reverse transcribed using Superscript II (Invitrogen) and an oligonucleotide with sequence complementarity to the 3’-linker as primer (6). Exonuclease I (USB) was added directly to the reverse transcription reaction and was incubated for 30 minutes at 37°C in 35mM Tris-HCl, 5mM (NH₄)₂SO₄, 42.5mM KCl, 1.5mM MgCl₂, 1mM MgSO₄, and 0.05% TritonX-100 to digest any lingering primers. A second 3’- linker ligation was performed as above using another 5’-adenylated, 3’-blocked oligonucleotide (5’ rAppCACTCGGGGACCAAGGA/ 3ddC/ 3’) followed by a third gel purification. The ligation product was amplified by PCR for 20 cycles with the following primers which contain sites recognized by the Ava I restriction enzyme: 5’ CACTCGGGGATGATGAGTCACAG 3’ and 5’ GTCCCTGTGCGCCGAGTG 3’.

The PCR product was diluted 100-fold and a second round of 10 cycles of PCR was performed. The PCR products were concatamerized by first digesting with Ava I, then
ligating the purified digestion products with T4 DNA ligase (New England Biolabs) overnight at room temperature. The entire ligation product was purified on a 4% NuSieve GTG Agarose gel (Cambrex) and concatamerized products >400 bp were extracted and treated with Taq polymerase which filled in and added nontemplated adenines to the 3’-ends. This was followed by cloning into the TOPO-TA vector (Invitrogen).

Upper bound estimates of ligatable 5’ termini in the pools of sel-1 related small RNAs associated with RNAi

To obtain an upper-bound estimate on the fraction of 5’-ligatable sel-1 siRNA among the total small sel-1-associated RNAs in the cell, we carried out the following calculation:

An estimate of the fraction of total small RNA that corresponds to sel-1 came from the 5’-ligation-independent cloning incidence (127/1612). An equivalent upper-bound estimate of 5’-monophosphate sel-1 small RNA (as a fraction of total small RNA) came from the 5’-ligation-dependent cloning incidence of 534/245420. These estimates would both miss any RNAs (both in the numerator and denominator) that were blocked on their 3’ ends. For the remainder of the following calculations, we will focus on only small RNAs that ligate on their 3’ end. As an additional basis for the calculation, we assume that other than the 5’-covalent structure, the two populations of small sel-1 RNA have roughly equivalent cloning efficiencies relative to the total. Under these assumptions, an upper bound on the fraction of sel-1 small RNAs with a ligation-available 5’-end (which we will designate "FSEL5") would be

\[
FSEL5 = \frac{\text{[sel-1 small RNAs with ligatable 5’-end]}}{\text{[total sel-1 small RNAs]}}
\]

This can be rewritten as

\[
FSEL5 = \frac{\{\text{[sel-1 small RNAs with ligatable 5’-end]}\}}{\{\text{[total sel-1 small RNAs]}\}}
\]

We next set F5=fraction of total small RNAs that have a ligatable 5’-end. We then have

\[
FSEL5 = \left\{ \frac{534/245420}{127/1612} \right\} \times F5 = 0.0276 \times F5
\]

Note that F5 is likely to be substantially less than 1, since many small RNAs in the cell appear to have modified 5’-ends. F5 values of less than one will give FSEL5<0.0276. From this calculation 2.76% (or about 1 in 40) becomes a rough upper bound for the fraction of sel-1 small RNAs capable of simple 5’ ligation. An indication that F5 might be <<1 comes from comparison of microRNA incidences in 5’ ligation-dependent and 5’-ligation independent cloning. Given differences of 5-10 fold in this incidence, it seems likely that FSEL5 is on the order of 1/200 to 1/400.
Upper limit estimates of primary siRNAs in the total sel-1 small RNA pool can be made based on assumptions that both strands are relatively equally represented in the primary siRNA pool with antisense heavily over-represented in the secondary siRNA pool. If this is the case, then one would estimate a maximum primary siRNA frequency among the 5'-ligation-dependent cloning experiments of \((2 \times \text{sense sel-1 siRNA}) \div (\text{total sel-1 siRNA})\).

Fifty one sense sequence clones were obtained from the 5'-ligation-dependent cloning, of which 47 were located in the trigger region. Twenty two of the latter sequences showed a \(\geq 9\) base complementarity to at least one antisense clone. For the following arguments, we have used the more conservative (upper bound) figure of 51 for the number of sense clones.

These assumptions give an estimate of primary siRNAs among the sequences from the 5'-ligation-dependent cloning experiments of \((2 \times 51)/534\) or 19.1%. Given an estimate that at most 2.76% of all sel-1 small RNAs are represented in the 5'-ligation-dependent library, we come up with a maximum fraction of primary sel-1 siRNAs among all sel-1 small RNAs induced during RNAi of \(0.191 \times 0.276 = 0.53\%\). Note that this could be 5-10 fold lower based on the value of F5 as discussed above. Based on similar arguments, we would estimate the incidence of primary sel-1 small RNAs among the total small RNA populations of the animal as not exceeding an upper limit of \(\sim 1/2500\) (0.042% = \(2 \times 51 \div 245420\)). This value reduces to \(\sim 1/25000\) if we assume that only 10% of the total population of small RNAs have ligatable 5'-ends.
Supporting text

Positional Analysis of small RNA structures associated with sel-1 silencing.

In each experiment, we observed a number of different incidences of small sel-1 associated RNAs which had identical sequences. Where such sequences were observed from a single experiment, there were several possibilities for their relationship.

1. Independent derivation

Given the size of the sel-1 coding region (2.4kb), there would be some chance of fortuitous coincidences where the same small RNA was recovered more than once. If all small RNAs were the same size and if all sites were present at equal abundance, we would expect a number of such coincidences among a large dataset such as the 454 sequences (for 534 sequences distributed uniformly in a segment of 2385 bp, we would expect approximately 60 such coincidences). Any bias in the incidence of observing small RNAs within the span of the gene would substantially increase this incidence. Thus the significant enrichment of siRNAs in the trigger homology and just upstream would be expected to dramatically increase the number of independent coincidences.

2. Duplication of a single cDNA clone during the initial PCR amplification.

The initial preparation of cDNA populations involves a PCR step which could amplify a single molecular instance of a given RNA so as to yield more than one clone and or sequence in later analysis.

3. Duplication of signals from amplification and pyrosequencing analysis from the 454-life-sciences platform.

Duplication of 454 signals can occur through inclusion of two beads in a single emulsion droplet during the 454-life-sciences sample preparation; in addition, optical contamination of a "blank" fiber optic element by an adjacent (filled and active) element cannot be ruled out. In a parallel experiment with independently derived DNA segments from C. elegans (7), this incidence was estimated to be approximately 10% of the total sequences recovered from a 454-life-sciences run.

Several features of DNA sequences that were recovered repeatedly suggest that the non-uniform distribution of small RNAs leading to "hot-spots" is a fundamental aspect of the RNAi response. This is most evident with several of the most frequently found small RNA species (Fig. 1). One segment of 22nt was recovered as a perfect and precise copy 22 times in the 454 dataset. This segment was in the center of the trigger region where siRNAs were most abundant. We observed instances in which the 454 sequencing primers were added to either end of the sequence, ruling out a single event amplified during the 454 reactions. An additional argument for some degree of independent derivation comes from the precise linker sequence junction with the cDNA. The joining of the 3' end of the initial siRNA to the adenylated linker during the first step of preparation for cDNA synthesis (6) produces a certain degree of heterogeneity in this junction. This could reflect some heterogeneity on the 5' end of the initial adenylated linker population or an as yet unidentified enzymatic activity acting on the 3'-termini of small RNAs as has been observed in plants (8). The fact that variant linker structures (fig. S2) are represented in the different sequences from this small segment of sel-1 argues for an independent derivation of at least some of these sequences. A third piece of evidence for bone-fide over-representation of this siRNA in the original population
comes from the observation of numerous slightly offset siRNAs from the same region (e.g., in addition to the total of twenty nine 22-mer instances at base 717, six 21-mer instances, and one 23-mer instance were derived from the same initiation point). These would be expected to derive from distinct siRNA molecules rather than duplication during cloning. Finally we note that several of the highly over-represented small RNAs were recovered multiple times in nearly equivalent forms (within one base) from both 5′-linker dependent and 5′-linker-independent cloning methods. Since these methods were carried out independently (and from distinct RNA preparations), it appears that such over-representation is not due to amplification of a single trigger molecule during cloning steps.

One mechanistically suggestive aspect of the distribution patterns emerges as the 5′-linker-dependent and 5′-linker independent patterns are compared: for several abundant small RNA species we see an offset in the precise 5′-position of the siRNA in that the 5′-ligation-dependent clones are often a single base offset (shorter) on their 5′-ends compared to the most abundant 5′-ligation-independent clones. This is shown in bulk in fig. S4. The graph in Figure S4 shows average offset values between arbitrary clones of each type, with a strong "mode" corresponding to a single base offset (with the 5′ end of the 5′-ligation-dependent clones most frequently corresponding to the second base of corresponding 5′-ligation-dependent clone). Examples at 669 and 717 (5′-termini) show such an effect for individual siRNAs. Note that some predominant small RNAs (e.g., 5′-termini at 777 and 794) that are present in the 5′-ligation-dependent clone population fail to appear at an enhanced frequency in the 5′-ligation-dependent population. Several conditions could account for this, including a very low efficiency of 5′ nucleotide removal (either in vivo or in vitro), unexpected sequence-dependent cloning biases in the two procedures, and the possibility that these small RNAs predominate in the primary (and not secondary) siRNA response.

The preferential occurrence of certain siRNA sequences in an ongoing response may reflect a combination of the initial processing of the bacterially derived trigger, an operational bias of the amplification mechanism, and/or preferential stabilization of certain effector siRNAs (potentially as a consequence of their high activity as surveillance guides). Whatever the source of this bias, the over-representation of certain specific sequences in the siRNA pool could lead to a situation in which a small number of over-represented siRNA epitopes could produce the type of off target effects observed by (9).
Figures

**Fig. S1.** FastA file containing all pyrosequenced clones obtained using 5’-ligation-dependent cloning of small RNAs from *C. elegans* undergoing RNAi for the *sel-1* gene. Small RNA sequence is bracketed by the 5’-linker sequence, 5’ ATCGTAGGCACCTGAAA 3’, and the 3’-linker sequence, 5’ CTGTAGGCACCACATCAAT 3’. The file can be accessed at www.sciencemag.org/cgi/content/full/1132839/DC1.
Fig. S2 Overrepresentation of a single 22-mer sequence in the 5'-ligation-dependent 454-life-sciences dataset: Evidence for bone-fide over-representation of this siRNA during the siRNA response. The following sequences are representative of the perfectly matched instances of these. See accompanying discussion for details.

Forward Orientation Instances
CCTACAGAGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
CCTACAGAGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
CCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
TGCCTACAGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
TGCCTACAGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
TGCCTACAGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
TGCCTACAGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC

Reverse Orientation Instances
CCTACAGCGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
GCCTACAGGAAACTGCACCTTTCCATATTATCTTTTCAGGTGC
Fig. S3 5’ ligation-independent small RNA cloning scheme. See Materials and Methods in Supplementary Material for details.

1. Size fractionate **small RNA** on 12% PAGE-urea
2. Ligate **5’-adenylylated linker #1** to 3’ end of small RNA
3. **Anneal primer** complementary to linker #1
4. **Reverse Transcribe**
5. **Denature**
6. Ligate **5’-adenylylated linker #2** to 3’ end of cDNA
7. **Gel purify**
8. PCR amplify with primers complementary to linkers 1 & 2
9. **Cut, concatemerize, clone, and sequence**
Fig. S4 The 5’-positions of sel-1 siRNA clones obtained using the 5’-ligation-dependent and independent cloning methods are offset by 1 nt. Graph shows the number of pairs of small RNA clones (one from 5’-ligation-dependent and one from 5’-ligation-independent libraries) which covered nearby regions of overlapping sequence and for which the 5’ end differed by the indicated number of base pairs. Thus the bar corresponding to an X axis value of 1 indicates the number of instances in which the ligation-dependent and ligation-independent clones covered an identical sequence region but with the ligation-independent clone a single base upstream of the ligation-dependent clone. Nucleotide positions corresponding to the 5’-termini of sel-1 siRNA clones compared between the two cloning methods revealed a peak of 1nt separation between the 5’-termini of clones from the 5’-ligation-dependent and –independent cloning methods.
**Fig. S5** Length comparison of *sel-1* siRNAs common to both 5’-ligation-dependent and independent cloning data sets. Pools of siRNAs generated during RNAi are heterogeneous in sequence (although there do appear to be “hotspots”; see fig. S4) and, thus, one can not assume that matching clones obtained by the two methods are indeed derived from the same siRNA. However, a cursory analysis shows that 76 clones from the 5’ ligation independent cloning match at least one sequence from the 5’-ligation-dependent method (193 different clones total). A distribution of the sizes of these clones show that those obtained from the 5’-ligation-independent method are 1nt longer.

![Size Distributions of common sel-1 siRNAs](image-url)
**Fig. S6** Examples of cloned small RNAs corresponding to exon/exon junction sequence of endogenous genes. Clones are from data set presented in fig. S1. Red text denotes cloned small RNA sequence, blue denotes intron sequence, black and green denote target mRNA sequence.

15__si2946_0.seq.1, C46C2.3  
1 clone, 5’ ligation-independent cloning

```
5’ ATTTCCGAAACGACGAAGACATCTGTGTCGTCGCCAGCACAATA 3’
3’ TGCTGCTTCTGTAGACAGCAGCCTT 5’
```

si2499.4, VW02B12L.3  
1 clone, 5’ ligation-independent cloning

```
5’ AGTTTTTTCATGTTTCTACAAGTTCTACAATGCGAAC 3’
3’ GTTACCAAGATGTTCGAAAC 5’
```

103589.2851.1378=56.1, T10A3.1a  
3 clones, 5’ ligation-dependent cloning

```
5’ TGAAAGTTTGTCCAGCTCCCTATGTAAAAGTGTATCT 3’
3’ ACAGGTCGAGGGATACATTTT 5’
```

013601.2190.3145=58.1,F59A3.3  
3 clones, 5’ ligation-dependent cloning

```
5’ GAACAGCCATGTGAAGCAAGATGGCAATTAAACGGTCCTG 3’
3’ TACACTTCGTTCACCGTTAATTT 5’
```

```
5’ GAACAGCCATGTGAAGCAAGATGGCAATTAAACGGTCCTG 3’
3’ TACACTTCGTTCACCGTTAATTT 5’
```
**Tables**

**Table S1** Small RNA profiles of various strains of *C. elegans*. Strains were subjected to RNAi by transgenic arrays expressing long hairpin RNAs (*unc-54*) or by feeding *sel-1* dsRNA-producing bacteria (*sel-1*) as indicated. Small RNAs were cloned using either a 5’-ligation-dependent method (6) or a 5’-ligation-independent method. miRNAs (blue text), miRNAs listed in miRBase (http://microrna.sanger.ac.uk/sequences/index.shtml); sequences were categorized as intergenic or coding (sequences antisense to exons highlighted in red text) based on annotation by UCSC Genome Bioinformatics; *sel-1/unc-54*, small RNAs corresponding to these sequences in animals undergoing RNAi for *sel-1* or *unc-54*, respectively, tncRNAs, tiny noncoding RNAs (10); SL2, splice leader 2; rRNA, ribosomal RNA; tRNA, transfer RNA; repeats, those sequences that matched more than one genomic locus; mitochondrial, sequence matching mitochondrial genome; other noncoding RNAs, two were found; one corresponding to *yrn-1* sequence, the other U5 snRNA. Yellow highlight, wild type (N2) strain; blue highlight, wild type (N2) animals fed on bacteria producing dsRNA corresponding to *sel-1* sequence. NCI, NGM with 25 µg/ml carbenicillin and 232.5 µg/ml IPTG. Numbers indicate percentages, numbers in brackets denote absolute values. Some RNA samples were treated with proteinase K (proteinase K), calf intestinal alkaline phosphatase (phosphatased) or heated to 65°C (65°C ) before cloning. Bound, RNAs were eluted from the 1st column during the miRvana miRNA isolation procedure (Ambion) and used in the cloning.
<table>
<thead>
<tr>
<th>Strain</th>
<th>Bacterial food</th>
<th>Growth condition</th>
<th>Stage of transfer</th>
<th>Treatment</th>
<th>Cloning Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>R14</td>
<td>asc-1 (phe-1)</td>
<td>NGM, 21°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R14</td>
<td>asc-1 (phe-1)</td>
<td>NGM, 21°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R17</td>
<td>asc-1 (phe-1)</td>
<td>NGM, 21°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R21</td>
<td>asc-1 (phe-1)</td>
<td>NGM+carb+IPTG, 16°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R21</td>
<td>asc-1 (phe-1)</td>
<td>NGM+carb+IPTG, 16°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R18</td>
<td>asc-1 (phe-1)</td>
<td>NGM+carb+IPTG, 16°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R25</td>
<td>asc-1 (phe-1)</td>
<td>NGM+carb+IPTG, 16°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R28</td>
<td>asc-1 (phe-1)</td>
<td>NGM+carb+IPTG, 16°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R36</td>
<td>asc-1 (phe-1)</td>
<td>NGM+carb+IPTG, 16°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R32</td>
<td>asc-1 (phe-1)</td>
<td>NGM+carb+IPTG, 16°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R43</td>
<td>asc-1 (phe-1)</td>
<td>NGM+carb+IPTG, 16°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R49</td>
<td>asc-1 (phe-1)</td>
<td>NGM+carb+IPTG, 16°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
<tr>
<td>R49</td>
<td>asc-1 (phe-1)</td>
<td>NGM+carb+IPTG, 16°C</td>
<td>young adults</td>
<td>mixed</td>
<td>proteinase K</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coding</th>
<th>3'UTR</th>
<th>5'UTR/exon</th>
<th>5'UTR</th>
<th>ex/intron</th>
<th>intron</th>
<th>exon</th>
<th>antisense</th>
<th>sense</th>
<th>sense</th>
<th>antisense</th>
<th>sense</th>
<th>antisense</th>
</tr>
</thead>
<tbody>
<tr>
<td>tRNA</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
</tr>
<tr>
<td>sel-1/unc-54</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>miRNAs</th>
<th>total</th>
<th>5' dependent</th>
<th>5' dependent</th>
<th>5' dependent</th>
<th>5' dependent</th>
<th>5' dependent</th>
<th>5' dependent</th>
<th>5' dependent</th>
<th>5' dependent</th>
<th>5' dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloning Method</td>
<td>proteinase K</td>
<td>Treatment</td>
<td>Bacterial food</td>
<td>Strain</td>
<td>RNA</td>
<td>3'UTR</td>
<td>5'UTR/exon</td>
<td>5'UTR</td>
<td>ex/intron</td>
<td>intron</td>
</tr>
<tr>
<td>proteinase K</td>
<td>mixed</td>
<td>mixed</td>
<td>mixed</td>
<td>mixed</td>
<td>mixed</td>
<td>mixed</td>
<td>young adults</td>
<td>young adults</td>
<td>young adults</td>
<td>mixed</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>siRNA</th>
<th>si813-908</th>
<th>si909-1004</th>
<th>si1005-1100</th>
<th>si1285-1380</th>
<th>si1381-1570</th>
<th>si1101-1188</th>
<th>si1955-1986</th>
<th>si1571-1602</th>
<th>si1667-1698</th>
<th>si1763-1858</th>
<th>si2627-2722</th>
<th>si2051-2146</th>
<th>si2147-2293</th>
<th>si2435-2530</th>
<th>si2723-3107</th>
</tr>
</thead>
<tbody>
<tr>
<td>tRNA</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td></td>
</tr>
<tr>
<td>sel-1/unc-54</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RNA</th>
<th>3'UTR</th>
<th>5'UTR/exon</th>
<th>5'UTR</th>
<th>ex/intron</th>
<th>intron</th>
<th>exon</th>
<th>antisense</th>
<th>sense</th>
<th>sense</th>
<th>antisense</th>
<th>sense</th>
<th>antisense</th>
</tr>
</thead>
<tbody>
<tr>
<td>tRNA</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
</tr>
<tr>
<td>sel-1/unc-54</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>siRNA</th>
<th>si813-908</th>
<th>si909-1004</th>
<th>si1005-1100</th>
<th>si1285-1380</th>
<th>si1381-1570</th>
<th>si1101-1188</th>
<th>si1955-1986</th>
<th>si1571-1602</th>
<th>si1667-1698</th>
<th>si1763-1858</th>
<th>si2627-2722</th>
<th>si2051-2146</th>
<th>si2147-2293</th>
<th>si2435-2530</th>
<th>si2723-3107</th>
</tr>
</thead>
</table>
Table S2 Quality and Accuracy of Sequences Obtained from 5’-ligation-dependent cloning. 106 clones obtained from the 5’-ligation-dependent cloning showed an imperfect match to sel-1 sequence. Below is a table summarizing the types of mutations observed. There appears to be a preponderance of single adenosine additions to the 5’ end of the siRNA. This bias is particularly prevalent among siRNAs corresponding to antisense sequence in the trigger homology region, although the paucity of clones outside the trigger region limits the significance of this comparison. If, however, this bias is truly reflective of a difference in the RNA pool, this trigger population, presumably enriched for primary siRNAs, may be the result of 5’-adenylation requiring a 5’-monophosphate. Mismatched clones from the 5’-ligation-independent method do not show this bias. Alternatively, we have not ruled out the possibility that the 5’-adenylation is merely a cloning artifact.

Table S2

<table>
<thead>
<tr>
<th></th>
<th>5’ ends</th>
<th>3’ ends</th>
<th>internal</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>T</td>
</tr>
<tr>
<td>upstream</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>up/trigger</td>
<td>30</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>trigger/down</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>downstream</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>antisense</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sense</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The pyrosequencing technology used for the 454-life-sciences high throughput sequencing has a high accuracy with the exception of regions of homopolymer sequence. Although C. elegans genome has a rather high homopolymer incidence, this is much more limited in the sel-1 mRNA sequence. The 2406 base sel-1 mRNA contains three A or T runs of 7 bases, three of 6 bases, and eight of 5 bases. No G or C runs of these lengths were present. Because of the short overall sequence lengths of siRNAs, our primary analysis was not designed to attempt assignments of sequences with insertions and deletions (the major error type in 454 sequencing). A small number of siRNAs corresponding to sel-1 may thus have been missed. Based on published quality analysis for 454 sequencing (11) and direct analysis of this dataset, the number of such clones that were "missed" is likely to be a minor fraction of the total. As an example, for the three 7 base homopolymer runs we observed a total of 6 perfectly matched clones, one putative clone which could have derived from sel-1 by a single base insertion and one that could have derived from a single base deletion. The single base insertion and deletion clones each corresponded to a sel-1 siRNA sequence for which a correct and precise sequence was obtained.
References